CIRCU

THE BULLETIN OF Mathematical BIOPHYSICS

UNIVERSITY OF [illegible]
LIBRARY
NOV 17 '49

SEPTEMBER 1949

A Note on the Two-Factor Theory with Rectification as Applied to Alternating Current Stimulation—*C. S. Patlak and H. D. Landahl* - - - - - - - - - - 149

Mathematical Biology of Social Behavior: II—*N. Rashevsky* 157

Input-Output Problems in Simple Nerve-Ganglion Systems *Alfonso Shimbel* - - - - - - - - - - - - - - 165

Some Suggestions for a Possible Approach to a Mathematical Biophysics of Mitosis—*N. Rashevsky* - - - - - 173

Outline of a Probabilistic Approach to Animal Sociology: I—*Anatol Rapoport* - - - - - - - - - - - - - - 183

Contribution to the Linkage Theory of Autopolyploids: II *Hilda Geiringer* - - - - - - - - - - - - - - 197

A Mechanism for Active Transport of Coenzyme-Like Substances with Possible Reference to Auxin Transport in Plants—*J. Boyd Best, Jr.* - - - - - - - - - - 221

Book Review—Joseph Schillinger, The Mathematical Basis of the Arts—*N. Rashevsky* - - - - - - - - - - 235

THE UNIVERSITY OF CHICAGO PRESS · CHICAGO · ILLINOIS

VOLUME 11 · · · · · NUMBER 3

THE BULLETIN OF MATHEMATICAL BIOPHYSICS

EDITED BY N. RASHEVSKY

The Bulletin is devoted to publications of research in Mathematical Biophysics, as described on the inside back cover.

THE BULLETIN is published by the University of Chicago at the University of Chicago Press, 5750 Ellis Avenue, Chicago, Illinois, quarterly, in March, June, September, December. ¶The subscription price is $4.00 per year, the price of single copies is $1.25. Orders for service of less than a full year will be charged at the single-copy rate. ¶Patrons are requested to make all remittances payable to The University of Chicago Press in postal or express money orders or bank drafts.

THE FOLLOWING are authorized agents:

For the British Empire, except North America, India, and Australasia: The Cambridge University Press, Bentley House, 200 Euston Road, London, N.W. 1. Prices of yearly subscriptions and of single copies may be had on application.

CLAIMS FOR MISSING NUMBERS should be made within the month following the regular month of publication. The publishers expect to supply missing numbers free only when losses have been sustained in transit, and when the reserve stock will permit.

BUSINESS CORRESPONDENCE should be addressed to The University of Chicago Press, Chicago, Ill.

COMMUNICATIONS FOR THE EDITOR and manuscripts should be addressed to N. Rashevsky, Editorial Office of the Bulletin of Mathematical Biophysics, 5741 Drexel Avenue, Chicago, Ill.

IMPORTANT ANNOUNCEMENT

The increase in the output of papers in the field of mathematical biophysics makes it difficult to insure prompt publication without an increase in the size of the journal. Therefore, the Bulletin of Mathematical Biophysics inaugurates the following service:

Upon acceptance of a paper, the Editor will, if necessary, ask the author to shorten the paper to an extent dictated by the requirements of a reasonably prompt publication. The shortening should in no case reduce the paper to a mere abstract. Such a shortened paper will be published within six months or less.

The unabbreviated original manuscript will be kept on file at the editorial office. Any person desiring to avail himself of the complete manuscript may obtain promptly a microfilm copy of the latter, at the cost of 2¢ per page plus a carrying charge of 25¢ and postage, by applying to the Editorial Office, 5741 Drexel Avenue, Chicago, Illinois.

All papers in the Bulletin which have been thus shortened will be marked at the end by the symbol MF followed by a figure indicating the number of double-spaced typewritten pages of the unabbreviated manuscript.

PRINTED BY THE DENTAN PRINTING COMPANY . . . COLORADO SPRINGS, COLORADO

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 11, 1949

A NOTE ON THE TWO-FACTOR THEORY WITH RECTIFICATION AS APPLIED TO ALTERNATING CURRENT STIMULATION

C. S. PATLAK AND H. D. LANDAHL
COMMITTEE ON MATHEMATICAL BIOLOGY
THE UNIVERSITY OF CHICAGO

The effect of rectification and non-normal accommodation on the two-factor theory of alternating current stimulation is studied. Rectification with normal accommodation increases the required amplitude of the current, especially for high frequencies. Rectification with non-normal accommodation introduces two upper limits, one for very low and one for very high frequencies.

The two-factor theories of N. Rashevsky (1948) and A. V. Hill (1933, 1936) account in a general way for the results of experiments with alternating currents. However if the principal time constant is determined independently, marked deviations occur at frequencies above 1000 cycles per second.

One of the possible causes for this deviation with which this paper will deal is rectification (Cole and Guttman, 1941). It was discovered that the squid giant nerve fiber will pass a greater amount of cathodal than anodal current for a given potential difference. The rectification ratio was found to be about 20:1. That is, one can say in this case that the electrical resistance of the nerve to an anodal current is 20 times greater than that for a cathodal current. For the frog sciatic nerve (Guttman, 1944) the rectification factor is about 1.03. The presence of a resistance due to the interstitial fluids makes it possible for the intrinsic rectification factor to be considerably larger. At the same time this resistance makes it possible for the nerve fibers to be under the influence of a partially rectified current even though the impressed current is of constant amplitude. This point will be discussed in a subsequent paper.

We will therefore break up the cycles into the positive half cycle (with a maximum amplitude of I_0) and a negative half cycle (with a minimum amplitude of αI_0 where α^{-1} equals the factor of rectification). Using Rashevsky's notation, we have

$$\left.\begin{aligned} d\varepsilon/dt &= KI_0 \sin\omega t - k(\varepsilon-\varepsilon_0) \\ dj/dt &= MI_0 \sin\omega t - m(j-j_0) \end{aligned}\right\} \quad I \geq 0;$$
$$\left.\begin{aligned} d\varepsilon/dt &= K\alpha I_0 \sin\omega t - k(\varepsilon-\varepsilon_0) \\ dj/dt &= M\alpha I_0 \sin\omega t - m(j-j_0) \end{aligned}\right\} \quad I \leq 0; \tag{1}$$

with $\varepsilon - j = 0$ as the condition for excitation. The angular frequency ω is equal to 2π times the frequency in cycles per second.

For convenience we will only give the solution of the equations for the ε's since the equations for the j's will be entirely analogous. These solutions are:

$$\begin{aligned} \varepsilon &= \varepsilon_0 + KI_0[k\sin\omega t - \omega\cos\omega t]/(\omega^2+k^2) + Ae^{-kt}, \quad I \geq 0; \\ \varepsilon &= \varepsilon_0 + K\alpha I_0[k\sin\omega t - \omega\cos\omega t]/(\omega^2+k^2) + A'e^{-kt}, \quad I \leq 0. \end{aligned} \tag{2}$$

Since at $t = 0$, $\varepsilon = \varepsilon_0$, therefore by solving for the constant for each half cycle, putting in the boundary value for the next half cycle and using the method of induction, we obtain, for $I \geq 0$, $2p\pi/\omega \leq t \leq (2p+1)\pi/\omega$,

$$\begin{aligned} \varepsilon &= \varepsilon_0 \\ &+ \frac{KI_0}{\omega^2+k^2}\left\{ k\sin\omega t - \omega\cos\omega t + \omega\left[(1-\alpha)\, e^{k\pi/\omega}\left(\frac{e^{2p\pi k/\omega}-1}{e^{k\pi/\omega}-1}\right)e^{-kt} \right] \right. \\ &\left. + \,\omega e^{-kt} \right\}, \end{aligned} \tag{3}$$

where the term within the brackets is obtained by the use of the formula for the sum of a geometric series. As there will be no excitation when $I < 0$, the solution for $I < 0$ is omitted.

Since we are only interested in steady state responses, we let $t \rightarrow \infty$, and therefore $p \rightarrow \infty$, in which case we obtain, for $I \geq 0$, $0 \leq t \leq \pi/\omega$,

$$\begin{aligned} \varepsilon &= \varepsilon_0 \\ &+ \frac{KI_0}{\omega^2+k^2}\left\{ k\sin\omega t - \omega\cos\omega t + \omega(1-\alpha)\frac{e^{k(\pi/\omega - t)}}{e^{k\pi/\omega}-1} \right\}. \end{aligned} \tag{4}$$

The corresponding expression for $I < 0$ is similar to this except that the sine and cosine terms are multiplied by α.

We shall not restrict ourselves to normal accommodation. For convenience we shall define β such that

$$M/m = \beta K/k. \tag{5}$$

We wish to find the current of minimum amplitude, $I_0 = I_m$, which will just produce excitation in the steady state. To find this value, which will be a function of the frequency and the parameters of the nerve, we first obtain $\varepsilon - j$ from equation (4) and its analogue and set it and its derivative equal to zero. The first of these equations involves $(j_0 - \varepsilon_0)$. This may be eliminated since the rheobasic current for a constant cathodal stimulation is equal to

$$I_R = k(j_0 - \varepsilon_0)/K\,[1 - \beta + \beta(1 - m/k)\,(\beta m/k)^{m/(k-m)}]\,. \tag{6}$$

We will therefore have to solve for the required I_m from the following two equations:

$$\begin{aligned} I_R[1 - \beta + \beta(1 - m/k)\,(\beta m/k)^{m/(k-m)}] \\ = I_m \left\{ \left[\frac{k^2}{\omega^2 + k^2} - \frac{\beta m^2}{\omega^2 + m^2} \right] \sin \omega t \right. \\ - \omega \left[\frac{k}{\omega^2 + k^2} - \frac{\beta m}{\omega^2 + m^2} \right] \cos \omega t + (1 - \alpha)\,\omega \left[\frac{k}{\omega^2 + k^2} \frac{e^{k(\pi/\omega - t)}}{e^{k\pi/\omega} - 1} \right. \\ \left. \left. - \frac{\beta m}{\omega^2 + m^2} \frac{e^{m(\pi/\omega - t)}}{e^{m\pi/\omega} - 1} \right] \right\}; \end{aligned} \tag{7}$$

$$\begin{aligned} 0 = \left[\frac{k^2}{\omega^2 + k^2} - \frac{\beta m^2}{\omega^2 + m^2} \right] \cos \omega t + \omega \left[\frac{k}{\omega^2 + k^2} - \frac{\beta m}{\omega^2 + m^2} \right] \sin \omega t \\ + (1 - \alpha) \left[\frac{\beta m^2}{\omega^2 + m^2} \frac{e^{m(\pi/\omega - t)}}{e^{m\pi/\omega} - 1} - \frac{k^2}{\omega^2 + k^2} \frac{e^{k(\pi/\omega - t)}}{e^{k\pi/\omega} - 1} \right]. \end{aligned} \tag{8}$$

The above equations can be solved explicitly only for very large or very small values of ω, the angular frequency. For very large ω, the above reduces to

$$\begin{aligned} I_m/I_R \\ = \frac{1 - \beta + \beta(1 - m/k)\,(\beta m/k)^{m/(k-m)}}{\{[2\sqrt{\pi^2 - (1 - \alpha^2)}\,k - (1 - \alpha)\pi\beta k]/2\pi\omega + (1 - \alpha)\,(1 - \beta)/\pi\}}, \end{aligned} \tag{9}$$

which approaches a constant. For $k = 2000$ and $m = 50\ \text{sec}^{-1}$, this constant is $0.9\pi/(1 - \alpha)\,(1 - \beta)$ if β is not too small.

For very small ω,

$$I_m/I_R = \frac{km[1 - \beta + \beta(1 - m/k)\,(\beta m/k)^{m/(k-m)}]}{\sqrt{\omega^2(m - \beta k)^2 + (1 - \beta)^2 k^2 m^2}}, \tag{10}$$

which also approaches a constant, $0.9/(1-\beta)$ for the same values of k and m as above.

Graphs have been plotted for some values of α and β. By the use of successive approximations, the value of t which would satisfy equation (8) for a given value of ω was calculated. This value of t was then substituted into equation (7) and the solution for I_m/I_R was found giving I_m/I_R as a function of ω.

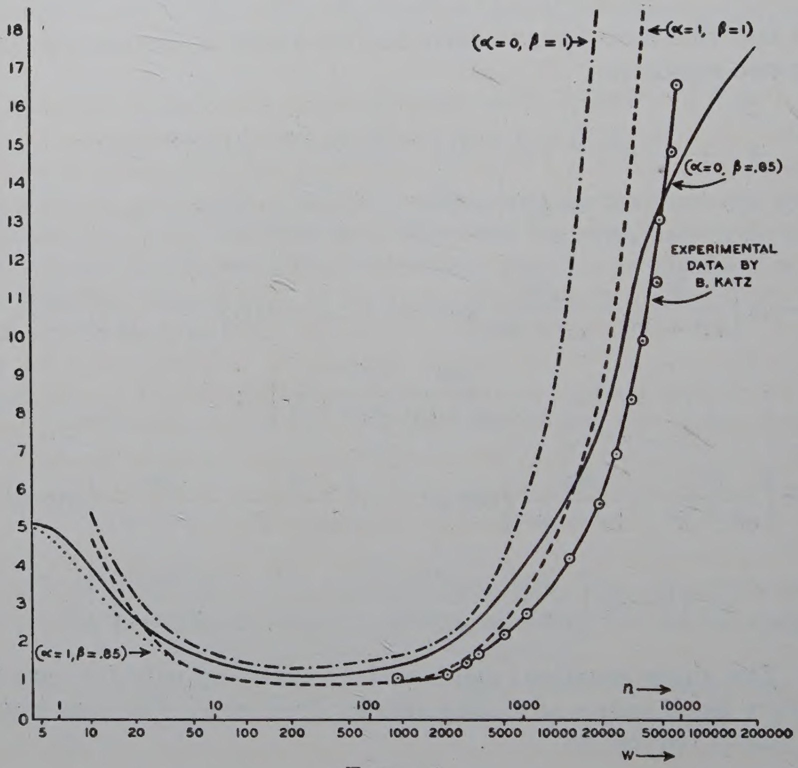


FIGURE 1

The experimental curve (solid, with points, Figure 1; Katz, 1939a, b) follows the simple two-factor theory equation (broken curve) for low frequency (Hill, Katz and Solandt, 1936), but for 500 cycles/sec. and above it deviates from this equation. From 5000 to 10,000 cycles/sec. the experimental curve follows the equation $I_m/I_R = \omega/2k$. It may be pointed out that in the experiments by B. Katz an amplitude modulated frequency technique permits one to disregard, to a large extent, the effect of a propagated impulse on the excitatory and inhibitory states. A value of $k = 2100$ sec^{-1} was determined by independent methods.

The curve (broken) for the two-factor theory with normal accommodation but without rectification ($\alpha = 1$, $\beta = 1$) as mentioned above, follows the experimental curve for low frequency, but for high frequency it is approximated by the equation $I_m/I_R = 0.9\ \omega/k$. For these as well as all other calculations k and m were taken to be 2000 and 50 sec^{-1} respectively.

The curve (upper) for the two-factor theory with complete rectification and normal accommodation ($\alpha = 0$, $\beta = 1$) deviates from the experimental curve at *all* points and at high frequency represents the equation $I_m/I_R = 1.8\ \omega/k$.

The curve (solid) for complete rectification but incomplete accommodation ($\alpha = 0$, $\beta = 0.85$) deviates from the experimental curve at most points, but not by so much that a suitable choice of α and β might not eliminate this deviation. The curve flattens out at a high frequency which was not experimentally reached. However, the curve bends markedly at $\omega = 5000$ whereas the data continue upward. The value of β was estimated from data by H. O. Parrack (1940).

The curve (dotted) for no rectification but incomplete accommodation ($\alpha = 1$, $\beta = 0.85$) is not appreciably different from the case for which $\alpha = 1$ and $\beta = 1$ except for low frequencies in which case it approaches the curve for which $\alpha = 0$, $\beta = 0.85$. The case in which $\alpha = 1$, $\beta \neq 1$ does permit the threshold value of the alternating current to be appreciably less than rheobasic (Householder, 1944) as is found in some experiments (Coppée, 1936).

Another test of the theory, besides that of carrying out experiments at very high and very low frequencies, is to excite with shocks at various intervals after the nerve has been stimulated with a subthreshold alternating current. In experiments of B. Katz at 5000 cycles/sec, an amplitude of 0.95 I_m, cut off and replaced by a direct current, gave regular responses at 0.85 I_R but few at $I_R/2$. According to the present simple formulation the required direct current, if the change was made at the most favorable point of the cycle, would be $0.05\ I_m = 0.4\ I_R$ since the experimental value of I_m/I_R is about 8 for this frequency.

If now a time τ is allowed to elapse with no current, then we find that if $\omega > 500$, $\tau > 3/k \approx 1.5$ m sec, $\theta \equiv I_0/I_R$, $I_0 < I_m$, and t^* is the time during the cycle at which the alternating current is shut off,

$$(\varepsilon - j)/(\varepsilon_0 - j_0) = 1 + \theta[\beta(1-\alpha)/\pi - \beta m \cos \omega t^*/\omega]e^{-m\tau}. \quad (11)$$

For high frequency ($\omega \gtrsim 4000$ per second) $(\varepsilon - j)/(\varepsilon_0 - j_0) - 1 = \theta\,\beta(1-\alpha)e^{-m\tau}/\pi$, a constant. That is, the amount of shock needed to excite will be greater than the rheobase I_R by a constant factor independent of frequency. If there is no rectification, the constant will be zero.

The effect of a local response (Hodgkins, 1938; Katz, 1939) has not been discussed in this paper. B. Katz (1939a, b) has suggested that this can account for the discrepancy at high frequencies. It may be noted, however, that since the time for disappearance of the local response is short, the values computed from equation (11) should agree with the experimental results if the theory is adequate.

It may be that the discrepancy at high frequencies can be accounted for by the theory of J. Th. Van der Werff (1948). However, he has given the results for the first cycle only, the result being the same as that given by the two-factor theories. What the result would be for the steady state is not clear because of the mathematical difficulties involved in the solution of the problem.

The introduction of rectification with normal accommodation does not improve agreement with experiment. However, rectification with non-normal accommodation does in some respects improve the agreement between theory and experiment over a considerable range of frequencies, though it introduces limits on the threshold current for which there is no evidence currently available.

This work was aided by a grant from the Dr. Wallace C. and Clara A. Abbott Memorial Fund of The University of Chicago.

LITERATURE

Cole, K. S. 1941. "Rectification and Inductance in the Squid Giant Axon." *J. Gen. Physiol.*, 25, 29–51.

Cole, K. S. and Guttman, R. 1941. "Electrical Rectification in Single Nerve Fibers." *Proc. Soc. Exp. Bio. and Med.*, 48, 293–297.

Coppée, G. 1936. "Stimulation by Alternating Current." *Cold Spring Harbor Symp.*, 4, 150–162.

Guttman, R. 1944. "Action of Potassium and Narcotics on Rectification in Nerve and Muscle." *Jour. Gen. Physiol.*, 28, 43–51.

Hill, A. V. 1936. "Excitation and Accommodation in Nerve." *Proc. Roy. Soc. London.* (*B*), 119, 305–355.

Hill, A. V., Katz, B., Solandt, D. V. 1936. "Nerve Excitation by Alternating Current." *Proc. Roy. Soc. London.* (*B*), 121, 74–133.

Hodgkins, A. L. 1938. "Subthreshold Potentials in a Crustacean Nerve Fiber." *Proc. Roy. Soc. London.* (*B*),126, 87–121.

Householder, A. S. 1944. "The Two-Factor Theory of Nervous Excitation with Non-Normal Accommodation." *Bull. Math. Biophysics*, 6, 157–161.

Katz, B. 1939a. *Electrical Excitation in Nerve.* London: Oxford University Press.

Katz, B. 1939b. "Nerve Excitation by High-Frequency Alternating Current." *Jour. Physiol.*, 96, 202–224.

Parrack, H. O. 1940. "Excitability of the Excised and Circulated Frog's Sciatic Nerve." *Amer. Jour. Physiol.*, 130, 481–495.

Rashevsky, N. 1948. *Mathematical Biophysics.* Second Edition. Chicago: University of Chicago Press.

Van der Werff, J. Th. 1948. *Biological Reactions caused by Electrical Currents and X-rays.* New York: Elsevier Pub. Co., Inc.

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 11, 1949

MATHEMATICAL BIOLOGY OF SOCIAL BEHAVIOR: II

N. RASHEVSKY
COMMITTEE ON MATHEMATICAL BIOLOGY
THE UNIVERSITY OF CHICAGO

A group of individuals is considered in which each individual has tendencies to exhibit one or another of two mutually exclusive behaviors. Neurobiophysically this may be described in terms of Landahl's reciprocally inhibited parallel reaction chains. The spontaneous excitations ε_1 and ε_2 at the central connections of each chain are a measure of the "natural" tendency of the individual toward one or the other of the two behaviors. According to equations derived by H. D. Landahl, the probability of one or the other behavior is determined by the difference $\varepsilon_1 - \varepsilon_2$. A population of individuals is considered in which $\varepsilon_1 - \varepsilon_2$ is distributed in some continuous way, and therefore in which the probability of a given behavior is distributed continuously between 0 and 1. The effect of other individuals exhibiting a given behavior is to increase the corresponding ε of the individual. Thus behavior of others affects the probability for a given behavior of each individual. It is shown that the equations describing the behavior of the population on the basis of this neurobiophysical picture reduce in the first approximation to the differential equations which were postulated by the author in his previous work on social behavior.

In a recent paper (Rashevsky, 1949, hereinafter referred to as *loc. cit.*) we outlined a neurobiological theory of social behavior based on the consideration of learning certain behavior patterns. We derived rather complicated expressions for the rate of change of the number of individuals from one behavior pattern to another.

In our recent book (Rashevsky, 1948a, hereinafter referred to as H. R.) we postulated some simple equations governing this rate of change, and we were able to apply those equations to a number of social phenomena. It is the purpose of the present paper to show how those formally postulated equations may be approximately derived from neurobiological considerations.

To this end we shall consider a somewhat simpler mechanism than in *loc. cit.* Instead of Landahl's "learning circuit," we shall consider Landahl's simple reciprocal inhibition circuit (Rashevsky, 1948b, chap. xxiv, hereinafter referred to as M. B.), which is reproduced on Figure 1 below.

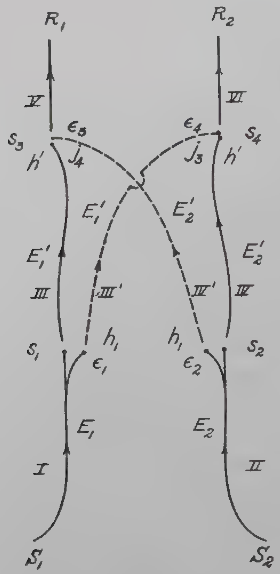


FIGURE 1

Let ε_1 and ε_2 be the amounts of ε factors which correspondingly excite the circuit so that reaction R_1 or reaction R_2 is produced. We shall interpret the connection excitations ε_1 and ε_2 as representing the desires of the individual for a reaction R_1 or R_2, respectively. The threshold h of the system (M. B. page 409) will be taken equal to zero. We shall denote the probability of reaction R_1 by P_1 and the probability of R_2 by P_2. By evaluating the integrals (7) and (8) on page 408 of M. B. we find

$$\text{for } \varepsilon_2 > \varepsilon_1; \quad P_1 = \tfrac{1}{2} e^{-k(\varepsilon_2 - \varepsilon_1)}; \quad P_2 = 1 - P_1; \tag{1}$$

$$\text{for } \varepsilon_1 > \varepsilon_2; \quad P_2 = \tfrac{1}{2} e^{-k(\varepsilon_1 - \varepsilon_2)}; \quad P_1 = 1 - P_2; \tag{2}$$

where k has the same meaning as in M. B.

An individual with $\varepsilon_1 = \varepsilon_2$, $P_1 = P_2 = \frac{1}{2}$, will have no preference for either reaction R_1 or R_2. If $\varepsilon_1 > \varepsilon_2$, then $P_1 > P_2$ and the individual shows a preference for R_1. As $\varepsilon_1 - \varepsilon_2$ tends to ∞, P_1 tends to 1. As $\varepsilon_2 - \varepsilon_1$ tends to ∞, P_2 tends to 1. Hence each individual may be characterized by his P_1, which varies from 0 to 1. Let $N(P_1)\,dP_1$ be the number of individuals having a P_1 between P_1 and $P_1 + dP$. If N_0 is the total number of individuals, we have

$$\int_0^1 N(P_1)\,dP_1 = N_0\,. \tag{3}$$

When an individual is in contact with another individual who exhibits behavior R_1, and when the latter influences the former to exhibit the same behavior, the situation results in an increased stimulation of the former individual toward activity R_1. That means that a quantity ϕ_1 is being added to ε_1 so that now we have $\varepsilon_1 + \phi_1 - \varepsilon_2$ instead of $\varepsilon_1 - \varepsilon_2$. Similarly, an influence toward performing reaction R_2 means an addition of a quantity ϕ_2 to ε_2. Unless ϕ_1 (resp. ϕ_2) is infinite an individual will always occasionally exhibit reaction R_2. The probability of the individual exhibiting reaction R_1 *now* (that is, under the influence of another individual who exhibits R_1) is, according to equation (2), given by

$$P_1' = 1 - \tfrac{1}{2}\, e^{-k(\varepsilon_1+\phi_1-\varepsilon_2)}\,, \tag{4}$$

and tends to 1 as ϕ_1 tends to infinity. It is largely a matter of social convention as to what probability for R_1 will be considered sufficiently large to say that the individual practically always shows reaction R_1. Let this arbitrary probability, which is close to 1, be denoted by p. In order to insure behavior R_1 we must have

$$1 - \tfrac{1}{2}\, e^{-k(\varepsilon_1+\phi_1-\varepsilon_2)} \geqslant p\,, \tag{5}$$

or

$$\phi_1 \geqslant \varepsilon_2 - \varepsilon_1 - \frac{1}{k}\log 2(1-p)\,. \tag{6}$$

Thus ϕ_1 must exceed a threshold value given by the right-hand side of (6) in order that this influence should insure behavior R_1. For $\varepsilon_1 - \varepsilon_2 > 0$ we have from (2)

$$\varepsilon_2 - \varepsilon_1 = \frac{1}{k}\log 2(1-P_1)\,. \tag{7}$$

For $\varepsilon_2 - \varepsilon_1 > 0$ we have from (1)

$$\varepsilon_2 - \varepsilon_1 = -\frac{1}{k}\log 2\,P_1\,. \tag{8}$$

Hence the threshold ϕ^* is equal to

$$\phi_1^* = \frac{1}{k}\log\frac{1-P_1}{1-p} \qquad \text{for } \varepsilon_1 > \varepsilon_2\,, \tag{9}$$

and to

$$\phi_1^* = -\frac{1}{k}\log 4\,P_1(1-p) \qquad \text{for } \varepsilon_2 > \varepsilon_1\,. \tag{10}$$

Thus ϕ_1^* is a function $\phi^*(P_1)$ of P_1. The same holds about ϕ_2^*, the threshold of ϕ_2.

When the additional strong enough stimulus ϕ_1 or ϕ_2 is applied, the corresponding reaction is not elicited immediately, but after a time τ, which is of the form

$$\tau = \tau_0 + \frac{1}{K_1} \log \frac{\phi_1}{\phi_1 - K_2}, \tag{11}$$

where the constants τ_0, K_1, and K_2 can be expressed in terms of the neurobiophysical constants of the neural pathways involved (M. B., chap. xxxiii). Thus it takes τ seconds for an individual to pass into the class of people exhibiting behavior R_1.

Very roughly we may set the rate of increase dx/dt of the number x of individuals which exhibit behavior R_1 as

$$\frac{dx}{dt} = \frac{1}{\tau}. \tag{12}$$

If τ_0 is very small and ϕ sufficiently large, then we have from (11) approximately

$$\tau = -\frac{1}{K_1} \log \left(1 - \frac{K_2}{\phi_1} \right) = \frac{K_2}{K_1 \phi_1} . \tag{13}$$

Equations (13) and (12) give

$$\frac{dx}{dt} = K \phi_1 ; \qquad K = \frac{K_1}{K_2}. \tag{14}$$

A similar relation holds between the rate dy/dt of increase of the number of individuals exhibiting behavior R_2 and the intensity of the influence ϕ_2.

The quantity ϕ_1 is itself a function of P_1 of the influencing individual. An individual influences another one toward a given behavior in the amount in which he himself exhibits this behavior. An individual with $P_1 = P_2 = \frac{1}{2}$, who when left to himself exhibits equally frequently both behaviors, will not influence others one way or another. We may, for example, set

$$\phi_1 = \kappa (P_1 - P_2) = \kappa (2P_1 - 1), \tag{15}$$

when κ is a purely sociological constant determined by the frequency of contacts between individuals, and where P_1 refers to the influencing individual. More generally we may put

$$\phi_1 = u_1 (P_1, \kappa). \tag{16}$$

Denote by $x(P_1)$ the number of individuals with a P_1 between P_1 and $P_1 + dP_1$, exhibiting behavior R_1. Then the total ϕ_1 in a given individual which results from the influence of all other individuals exhibiting behavior R_1, is given by

$$\phi_1 = \int_0^1 u_1(P_1, \kappa)\, x(P_1)\, dP_1 . \tag{17}$$

Similar considerations apply to ϕ_2, which measures the influence toward behavior R_2. Denoting by $y(P_1)$ the number of individuals with a P_1 between P_1 and $P_1 + dP_1$ exhibiting behavior R_2, and by $u_2(P_1, \kappa)$ a function corresponding to u_1, we have

$$\phi_2 = \int_0^1 u_2(P_1, \kappa)\, y(P_1)\, dP_1 . \tag{18}$$

We also have

$$x(P_1) + y(P_1) = N(P_1). \tag{19}$$

From (14), (17), and (18) we have

$$\frac{dx(P_1)}{dt} = K \int_0^1 u_1(P_1, \kappa)\, x(P_1)\, dP_1 - K \int_0^1 u_2(P_1, \kappa)\, y(P_1)\, dP_1 . \tag{20}$$

An individual for whom

$$\phi_1^* > \int_0^1 u_1(P_1, \kappa)\, N(P_1)\, dP_1 \tag{21}$$

will for practical purposes never exhibit behavior R_1, because even if all the other individuals would exert an influence on him in that direction, that influence, as given by the right-hand side of (21), would not be sufficient. Hence

$$\phi_1^*(P_1) = \int_0^1 u_1(P_1, \kappa)\, N(P_1)\, dP_1 \tag{22}$$

gives us an equation for the determination of the quantity P_1 such that all individuals having a P_1 above that quantity cannot be induced to perform reaction R_2. Let us denote that quantity by $1 - \Delta_1$. In a similar fashion we find a quantity Δ_2 such that all individuals having a P_1 between 0 and Δ_2 cannot be induced to perform reaction R_1. In the terminology of H. R. the two types of individuals would be called actives of type I and actives of type II. The total number of actives of type I is given by

$$X_0 = \int_{\Delta_1}^1 N(P_1)\, dP_1 . \tag{23}$$

The total number of actives of type II is given by

$$Y_0 = \int_0^{\Delta_2} N(P_1)\,dP_1\,. \tag{24}$$

All other individuals are, in that terminology, passive ones. The number of passive individuals exhibiting behavior R_1 is

$$X = \int_{\Delta_2}^{1-\Delta_1} x(P_1)\,dP_1\,, \tag{25}$$

while the number of passives exhibiting behavior R_2 is given by

$$Y = \int_{\Delta_2}^{1-\Delta_1} y(P_1)\,dP_1\,. \tag{26}$$

Since the right-hand side of (20) is a constant independent of P_1 we may integrate both sides with respect to P_1, between the limits 0 and 1, remembering that X_0 does not change with respect to time and therefore that $d(X_0 + X)/dt = dX/dt$. We find because of (24) and (25)

$$\frac{dX}{dt} = K\int_0^1 u_1(P_1\,,\kappa)\,x(P_1)\,dP_1 - K\int_0^1 u_2(P_2\,,\kappa)\,y(P_1)\,dP_1\,. \tag{27}$$

Now put

$\bar{a}_0 =$ an appropriate average of $u_1(P_1)$ in the interval $(1-\Delta_1\,, 1)$;

$\bar{a} =$ an appropriate average of $u_1(P_1)$ in the interval $(\Delta_2\,, 1-\Delta_1)$;

(28)

$\bar{c}_0 =$ an appropriate average of $u_2(P_1)$ in the interval $(0\,, \Delta_2)$;

$\bar{c} =$ an appropriate average of $u_2(P_1)$ in the interval $(\Delta_2\,, 1-\Delta_1)$;

and denote

$$a_0 = K\,\bar{a}_0\,; \qquad a = K\,\bar{a}\,; \qquad c_0 = K\,\bar{c}_0\,; \qquad c = K\,\bar{c}\,. \tag{29}$$

Using the mean value theorem we may now write equation (27)

$$\frac{dX}{dt} = a_0X_0 + aX - cY - c_0Y_0\,. \tag{30}$$

Equation (2) of chapter iii of H. R. is a special case of (30) if we make $a = c$ in the latter.

It is, however, important to remark that the coefficients a_0 , a , c , and c_0 in equation (30) are in general functions of time, because $x(P_1)$ and $y(P_1)$ contain time as a parameter. On the other hand, in equation (2) of chapter iii of H. R., the coefficients are assumed to be constant. If, however, as a crude approximation we substitute for the coefficients a_0 , a , c_0 and c in equation (30) some appropriate time averages, then equation (2) of chapter iii of H. R. is obtained as a sort of zeroth approximation to equation (30) and hence to equation (27).

Thus we have in principle expressed the coefficients of influence, used previously, the terms of the neurobiophysical parameters, K_1 , K_2 , k , and of the sociological constant κ .

LITERATURE

Rashevsky, N. 1948a. *Mathematical Theory of Human Relations.* Bloomington: The Principia Press.

Rashevsky, N. 1948b. *Mathematical Biophysics.* Revised Edition. Chicago: University of Chicago Press.

Rashevsky, N. 1949. "Mathematical Biology of Social Behavior: I." *Bull. Math. Biophysics,* 11, 105–113.

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 11, 1949

INPUT-OUTPUT PROBLEMS IN SIMPLE NERVE-GANGLION SYSTEMS

Alfonso Shimbel
Committee on Mathematical Biology
The University of Chicago

The dynamics of simple nerve-ganglion systems is studied. The frequency of nerve impulses (action currents) leaving the ganglion is expressed as a function of the frequency of impulses entering the ganglion. The problem is varied by considering ganglia having internal association connections.

Consider a nerve (Figure 1) containing N axones which "map into" a ganglion consisting of an equal number of cell bodies. The axones of the ganglionic cells, each receiving a connection from the "incoming bundle" form the "outgoing bundle."

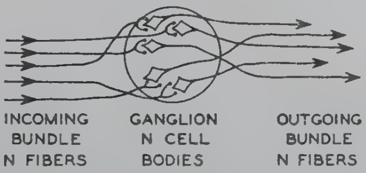


Figure 1

At the outset it will be necessary to introduce a parameter referred to as the "firing frequency" which has the following meaning. Arbitrarily choose a "small" interval of time δ (say from $t - \frac{1}{2}\delta$ to $t + \frac{1}{2}\delta$), then the average number of "action currents" occurring in a fibre bundle during the chosen interval of time *per unit time* per axone will be called the *firing frequency* of that bundle at the time t.

Now let $\phi(t)$ (the input) and $f(t)$ (the output) denote the firing frequencies of the incoming and outgoing bundles respectively. In what follows we will investigate the relation between the input and output of a system such as the one described above.

To begin with very little can be said about the relation between $\phi(t)$ and $f(t)$ unless we have some knowledge about the nature of the connections found in the ganglion.

For some of the simpler cases (e.g., the case in which the connections are random and the duration of the synaptic delay is the same for each fiber) the relation between $\phi(t)$ and $f(t)$ is rather evident. In the example cited above $f(t)$ is equal to the value which $\phi(t)$ had at a time "one synaptic delay" σ in the past, thus:

$$f(t)=\phi(t-\sigma). \tag{1}$$

In general, however, we can expect that the synaptic delays will vary from cell to cell and will also depend upon the physiological environment of the cell as a function of time. In order to include the more general case of a ganglion which is characterized not by one synaptic delay but by a distribution of them, we will introduce the notation $K(\sigma)\,d\sigma$ which denotes the probability that an arbitrarily chosen neuron is characterized by a synaptic delay in the range $\sigma-\frac{1}{2}d\sigma$ to $\sigma+\frac{1}{2}d\sigma$.

Under these circumstances we can expect that $f(t)$ will be some transformation of $\phi(t)$ which depends upon the form of $K(\sigma)$. More specifically, since $\phi(t-\sigma)$ is the firing frequency of the incoming bundle at the time $t-\sigma$, the product $\phi(t-\sigma)K(\sigma)\,d\sigma$ is the fraction of $f(t)$ which is due to the input at the time $t-\sigma$. It follows then that if $\phi(t)=0$ for the time interval $t=-\infty$ to $t=0$ that

$$f(t)=\int_0^t K(\sigma)\phi(t-\sigma)\,d\sigma. \tag{2}$$

By a simple transformation of variables this equation can be transformed into a more familiar form. For this purpose let $\sigma=t-\tau$ whereupon $d\sigma=-d\tau$. It follows then that when $\sigma=0$, $t=\tau$ and when $\sigma=t$, $\tau=0$. Equation (1) then becomes

$$f(t)=\int_0^t K(t-\tau)\phi(\tau)\,d\tau. \tag{3}$$

Of course very little more can be said about the relation between input and output than is expressed in equation (2) until more is known about the form of the function $K(\sigma)$. In any study of actual nerve-ganglion systems the form of $K(\sigma)$ would need to be determined empirically. For the purpose of our analysis, however, we will assume a form of $K(\sigma)$ in order to determine the "kinds" of relations which such a system implies. For this purpose let $K(\sigma)=Ke^{-K\sigma}$ where K is a constant. Inserting this form of $K(\sigma)$ into equation (2) we obtain

$$f(t)=Ke^{-Kt}\int_0^t e^{K\tau}\phi(\tau)\,d\tau. \tag{4}$$

The integral of equation (3) can be formally evaluated if the operation of integration by parts is performed an indefinite number of times.

Using this device, equation (3) after n integrations can be written as

$$f(t) = \phi(t) - \frac{1}{K}\phi'(t) + \frac{1}{K^2}\phi''(t) - \cdots + \frac{(-1)^{n-1}}{K^{n-1}}\phi^{(n-1)}(t)$$
$$- e^{-Kt}\left[\phi(0) - \frac{1}{K}\phi'(0) + \frac{1}{K^2}\phi''(0) \right. \tag{5}$$
$$\left. - \cdots + \frac{(-1)^{n-1}}{K^{n-1}}\phi^{(n-1)}(0)\right] + R_n,$$

where $R_n = (-1)^n K e^{-Kt} \int_0^t \frac{1}{K^n} e^{K\tau}\phi^{(n)}(\tau)\,d\tau.$

Equation (5) implies that the output $f(t)$ is no longer simply a temporal translation of the input as was the case for a constant synaptic delay. On the contrary, when the synaptic delays are distributed as indicated above, the output may be a serious *distortion* of the input.

It follows from equation (5) that for $t >> 0$ we can write approximately

$$f(t) = \phi(t) - \frac{1}{K}\phi'(t) + \frac{1}{K^2}\phi''(t) - \cdots + \frac{(-1)^{n-1}}{K^{n-1}}\phi^{(n-1)}(t) + R_n. \tag{6}$$

It can be seen from equation (6) that if $\phi(t)$ is a constant then after a sufficiently long time $f(t) = \phi(t)$.

Note, however, that if $\phi(t)$ is analytic and $\bar{\sigma}$ a constant sufficiently small, we can write by means of Taylor's theorem

$$\phi(t - \bar{\sigma}) = \phi(t) - \bar{\sigma}\phi'(t) + \frac{\bar{\sigma}^2}{2!}\phi''(t) - \cdots + R_n. \tag{7}$$

By comparing equations (6) and (7), it is evident that if $\phi(t)$ is a linear function of time, i.e., if all derivatives after the first are zero then

$$f(t) = \phi(t - \bar{\sigma}). \tag{8}$$

In other words, if the input is a linear function of time, then after a sufficiently long time the output will be the same linear function with a *time lag* of $\bar{\sigma} = \frac{1}{K}$. It is interesting to note at this point that $\bar{\sigma} = \frac{1}{K}$ is the *average* synaptic delay of all the cells in the ganglion.

For polynomials of degree greater than one we might well expect more serious distortions. In fact for a parabolic input, the output is distorted by a constant, and in general for an input represented by a polynomial of degree n, the output is distorted by a polynomial of degree $n - 2$.

There are a great many important factors found in actual nervous systems which are completely neglected by this model. Nevertheless it is interesting to note, if only in passing, that the distortions of input which were discussed above are suggestive of such phenomena as *adaptation* and *facilitation.*

Consider now the case in which the ganglion neurons not only receive connections from the incoming bundle but also have an "association mapping." In other words each neuron in the ganglion receives one connection from another ganglion cell and also "sends out" one connection to another ganglion cell.

In this case the output $f(t)$ is not due simply to $\phi(t)$ and $K(\sigma)$ but also is influenced by the self-mapping.

Before considering the general equation governing the relation between the input and output let us examine the implications of a system of three neurons connected in the manner depicted by Figure 2.

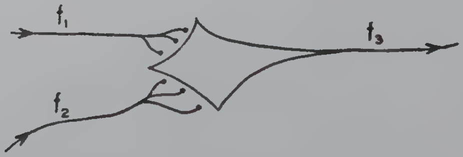


FIGURE 2

Let f_1, f_2 and f_3 be the firing frequencies of the respective neurons, and assume that the output neuron will fire if either neuron 1 or neuron 2 or both 1 and 2 fire. The expectancy of firing of neuron 1 for a picked interval of time δ is given by δf_1 and similarly that of neuron 2 is δf_2. Now since the probability that 1 and 2 will fire at exactly the same time is zero, it follows that if the refractory period of each neuron is also zero then

$$\delta f_3 = \delta f_1 + \delta f_2 , \tag{9}$$

whereupon we obtain

$$f_3 = f_1 + f_2 \, . \tag{10}$$

Equation (10) which tells us that the output frequency is simply the sum of the two input frequencies can be considered valid for those systems in which the refractory period is vanishingly small. We know however that this is usually not the case. Therefore, in order to account for those times when neurons 1 and 2 both fire within a period of time less than or equal to the refractory period of neuron 3 (and therefore are counted only as a *single* impulse) , equation (9) must be modified to read

$$\delta f_3 = \delta f_1 + \delta f_2 - \delta^2 f_1 f_2 \, , \tag{11}$$

where now the symbol δ refers to the magnitude of the refractory period. It follows then that

$$f_3 = f_1 + f_2 - \delta f_1 f_2 \, . \tag{12}$$

The output frequency f_3 is seen then to be a function not only of the input frequencies but also of the refractory period δ .

Returning now to the nerve-ganglion model with the additional assumption that the ganglion has a "self-mapping" we find that each outgoing axone is a member of a three neuron system such as the one just described. In this model f_1 corresponds to the input frequency and f_2 corresponds to the "self-mapping" or output frequency.

Reasoning as before but this time taking into account the self-mapping we obtain

$$\begin{aligned} f(t) = {} & \int_0^\infty K(\sigma)\phi(t-\sigma)\,d\sigma + \int_0^\infty K(\sigma) f(t-\sigma)\,d\sigma \\ & - \delta\left[\int_0^\infty K(\sigma)\phi(t-\sigma)\,d\sigma\right]\left[\int_0^\infty K(\sigma) f(t-\sigma)\,d\sigma\right] . \end{aligned} \tag{13}$$

In formulating equation (13) it was assumed that the self-mapping was complete, i.e., each member of the ganglion possesses a self-mapping axone. If we change this requirement to the more general case of a ganglion having only a fraction λ of its neurons in a self-mapping configuration we must modify equation (13) slightly. For this case we have

$$\begin{aligned} f(t) = {} & \int_0^\infty K(\sigma)\phi(t-\sigma)\,d\sigma + \lambda \int_0^\infty K(\sigma) f(t-\sigma)\,d\sigma \\ & - \delta\,\lambda\left[\int_0^\infty K(\sigma)\phi(t-\sigma)\,d\sigma\right]\left[\int_0^\infty K(\sigma) f(t-\sigma)\,d\sigma\right] . \end{aligned} \tag{14}$$

Now if we consider as previously that $K(\sigma) = Ke^{-K\sigma}$ and $\phi(t) = 0$ for $-\infty < t \leq 0$, and again as before if we put $t - \sigma = \tau$, then equation (14) can be written as

$$f(t) = Ke^{-Kt} \int_0^t \phi(\tau) e^{K\tau} d\tau + \lambda Ke^{-Kt} \int_0^t f(\tau) e^{K\tau} d\tau$$
$$- \delta \lambda \left[Ke^{-Kt} \int_0^t \phi(\tau) e^{K\tau} d\tau \right] \left[Ke^{-Kt} \int_0^t f(\tau) e^{K\tau} d\tau \right]. \tag{15}$$

For certain simple forms of the input $\phi(t)$, equation (15) can be readily solved.

We will here consider only the case of constant input, i.e., where $\phi(t) = \phi_0 =$ constant. Furthermore, only the asymptotic behavior of $f(t)$ will be studied, that is, $f(t)$ will be examined for large values of t.

By differentiating equation (15) and letting

$$I \equiv Ke^{-Kt} \int_0^t f(\tau) e^{K\tau} d\tau$$

we obtain

$$f'(t) = K\phi_0 e^{-Kt}$$
$$+ \lambda(1 - \phi_0 \delta + \phi_0 \delta e^{-Kt}) I' - \lambda K \phi_0 \delta e^{-Kt} I. \tag{16}$$

However from the definition of I it can be seen that

$$I' = Kf(t) - KI. \tag{17}$$

By solving equation (15) for I and substituting this form into equations (16) and (17) we obtain a system of two equations with two unknowns. By simultaneous solution we then obtain $f'(t)$ as a function of t, ϕ_0, K, δ, λ, and $f(t)$.

By considering the structure of the nerve-ganglion model we can infer that for a constant input, $f(t)$ must be a monotonically increasing function. However due to the finite refractory periods of the neurons, $f(t)$ has an upper limit of $1/\delta$. From this we can conclude that for a sufficiently large t, $f(t)$ is approximately constant. In other words $f'(\infty) = 0$.

Using this information it becomes possible to solve for $f(\infty)$, and obtain

$$f(\infty) = \frac{\phi_0}{1 - \lambda(1 - \delta \phi_0)}. \tag{18}$$

Note that for $\lambda = 1$, that is, for the case of a nerve-ganglion system with a complete self-mapping, equation (18) reduces to $f(\infty) = 1/\delta$ quite regardless of the magnitude of the constant input of ϕ_0. This latter result, however, can be immediately predicted by examining the model under consideration when $\lambda = 1$. For this case *every* ganglionic cell has a self-mapping axone so that no matter how small the constant input ϕ_0 may be, if we wait sufficiently long all of them will eventually be firing at their maximum frequency. As was pointed out before, this maximum frequency is limited by the finite magnitude of the refractory period and equals the value $1/\delta$.

Note also that for the other extreme value of λ, namely, $\lambda = 0$, the solution is again trivial, that is, $f(\infty) = \phi_0$. In other words, if the ganglion cells have no self-mapping, then the input goes through the ganglion undistorted.

In general, however, for all λ such that $0 < \lambda < 1$ we must refer to equation (18).

The author is indebted to Mr. George Karreman for a critical reading of this paper.

This work was aided by a grant from the the Dr. Wallace C. and Clara A. Abbott Memorial Fund of The University of Chicago.

LITERATURE

Rapoport, A. 1948. "Cycle Distributions in Random Nets." *Bull. Math. Biophysics,* **10**, 145–157.

Rapoport, A. and Shimbel, A. 1948. "Steady States in Random Nets: I." *Bull. Math. Biophysics,* **10**, 211.

Rapoport, A. 1948. "Steady States in Random Nets: II." *Bull. Math. Biophysics,* **10**, 221.

Shimbel, A. and Rapoport, A. 1948. "A Statistical Approach to the Theory of the Central Nervous System." *Bull. Math. Biophysics,* **10**, 41-55.

Shimbel, A. 1948. "An Analysis of Theoretical Systems of Differentiating Nervous Tissue." *Bull. Math. Biophysics,* **10**, 131–143.

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 11, 1949

SOME SUGGESTIONS FOR A POSSIBLE APPROACH TO A MATHEMATICAL BIOPHYSICS OF MITOSIS

N. RASHEVSKY
COMMITTEE ON MATHEMATICAL BIOLOGY
THE UNIVERSITY OF CHICAGO

A tentative suggestion is made to interpret the phenomenon of mitosis as being due to coiling and uncoiling of long chain molecules, which are assumed to constitute the spindle. A possible quantitative approach is briefly outlined.

The theory of metabolic diffusion forces has proven to be rather useful in the theory of the general phenomenon of cell division, and has led to correct predictions of the over-all quantitative features of this process (Landahl, 1942; Rashevsky, 1948). While the present developments do not by any means prove that the diffusion forces are the only, or even the principal, cause of cell division, they demonstrate definitely that in many instances they are a sufficient cause. Moreover the correct prediction of the time dependence of the elongation and constriction of freely dividing cells shows that even if other factors should eventually be found to be the cause of cell division, any new theory would have to lead to the same, or to very similar, equations for elongation and constriction.

The situation is, however, quite unsatisfactory when we consider other important special aspects of cell division such as mitosis. There is no "typical" mitosis any more than there is a "typical" cell division. The process of mitosis in different types of cells differs considerably in details. One general feature, however, which is common to all cases stands out: Mitosis provides a mechanism for an equal distribution of chromosomes in the daughter-cells. The arrangement of chromosomes in the metaphase in an equatorial plate and the subsequent movement of each set of daughter-chromosomes toward the opposite poles of the achromatic spindle is a general feature common to all mitoses. The greater complexity of the whole phenomenon, as compared to the relative simplicity of the over-all process of cell division, makes a simple explanation rather unlikely. The more complex an explanation we may have to use, the more numerous will be such explanations. The only proper way of procedure would be the one used previously, namely, to investigate all, or as many as possible, theoretically conceivable explanations *in abstracto*, then study the mathematical consequences of all those approaches, and finally, from comparison of those consequences with observation, decide which of the approaches is the most useful.

Whatever the nature of the chromosomes, if their arrangement in an equatorial plate should represent, at least temporarily, a stable configuration, there are only two conceivable types of forces that would produce such a configuration. Either each chromosome must be repelled by the poles with a force which *decreases with the distance* from the center of repulsion, or each chromosome must be attracted toward each pole by a force which *increases with the distance* from the center of attraction. If the poles are metabolic centers *producing* some metabolite S the first type of forces could be interpreted as diffusion forces. If the chromosomes are attached to each pole by contractile fibers whose force of contraction increases with their extension, we will have the second type of forces. A combination of both forces was suggested in an outline of a mathematical theory of mitosis by N. Rashevsky (1938). Some quantitative consequences were drawn from the assumption that the daughter-chromosomes are pulled apart by contractile fibers, and were compared to the meager data available (Rashevsky, 1941). As F. Schrader (1944) pointed out, however, the nature of the data used was such as to make the apparent agreement meaningless.

The shortcomings of the approach outlined above do not yet necessarily preclude its usefulness if it is developed further and improved. In the present paper we shall, however, outline a different possible approach.

It is generally accepted that the chromosomes are aggregates of molecules which perform important catalytic functions in the life of the cell. In a sense each of them may be viewed as a giant molecule which consists of complex parts, the genes. Argument about terminology would be rather fruitless since we may refer to a polymerized molecule as one giant molecule or as an aggregate of smaller ones.

The circumstance that usually the chromosomes are not visible in the resting nucleus may be interpreted as follows. The chromosome may be viewed as a long chain molecule whose cross dimensions are too small to make it visible. During the process of mitosis the action of certain chemicals, or perhaps the change in pH , makes each chromosome the center of precipitation of chromatin and thus makes it visible. The situation is not unlike making individual ions visible in a Wilson cloud-chamber.

Though not conclusive there is evidence that each chromosome is attached at its spindle attachment to a fiber which forms a part of the achromatic spindle. Let us assume that what we observe as a chromosome is only the temporarily visible part of a T-shaped molecule, which we shall denote as a C–molecule, as shown in Figure 1a. The chromosome proper we shall designate as the "c–part" of the C–molecule. The other part of it we shall call the f–part. Let all C–molecules have f–parts of the same or approximately the same length.

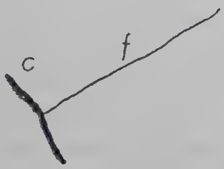


FIGURE 1a

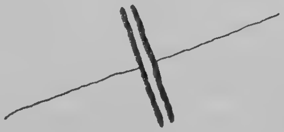


FIGURE 1b

During the life cycle of the cell such a molecule is duplicated by the formation of a sort of "mirror image" (Figure 1b) and then split into two daughter molecules. It is a problem of molecular physics to develop a theory of such self-duplicating molecules.

As a catalyst each chromosome is responsible for the formation of some substance. Naturally the production of each substance will depend on the particular "life-phase" of the chromosome. Some of these substances, S, may be produced mostly during the period following the formation of the daughter chromosomes. During the time just preceding the separation of the daughter chromosomes when the C–molecules are in configuration, shown in Figure 1b, the production of S will be at a minimum.

Due to the thermal agitation the C–molecules will be "knotted"

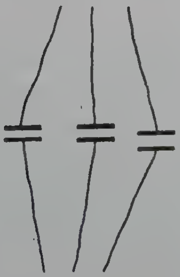


FIGURE 2a

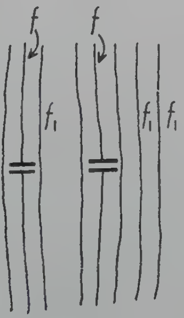


FIGURE 2b

at random as discussed by W. Kuhn (1934, 1936). Suppose that at the time when all C–molecules have the configuration shown on Figure 1b something causes their f–parts to orient themselves parallel to each other. A configuration shown in Figure 2a will then result which reminds us of the general shape of a metaphase nucleus. The reason for such an orientation may be either an external force or a sort of "inner molecular field," which may appear at a proper time. The appearance of such an orienting field may be connected with the presence or absence of some metabolites. For instance, the presence of S may cause the weakening or complete abolition of certain chemical bonds which tend to hold the f–parts parallel to each other. In that case when the S-substance is at a minimum, these bonds will come into play and orient the f–parts. Simple quantitative estimates which we shall discuss presently indicate, however, that with plausible bond energies there can be no self-orienting effect of the f–parts, if the total number of C–molecules is small, that is, of the order of ten or a hundred. If, however, we assume the presence of numerous long molecules in the nucleus which are not all attached to the chromosomes (f_1–molecules), then the decrease in a certain S–substance may well cause the orientation of the f_1– molecules. This orientation creates an "inner molecular field" which also orients the f–parts of the C–molecules. Thus the picture shown in Figure 2b results. Now if the c–parts of the C– molecules separate, either due to the tension of the f–parts or for other reasons, this will result in an increased production of S, which will loosen the bonds between the f_1–molecules and the f–parts and result in the latter becoming "knotted" again. This knotting results in a contraction of the f–chains, and thus pulls the c–parts, or daughter chromosomes, away.

If the f_1–chains are also susceptible to the action of S, but to a different degree so that they break in the middle *after* the chromosomes have been pulled apart toward their respective poles, then the f_1–chains also will break and coil up, forming the daughter nuclei with the halves of the f–chains.

We may make, if we wish, a more specific assumption. Let the f– and f_1–chains consist of ionized molecules, the ionization depending on the pH of the medium. The pH in this case may play the role of the metabolite, S. At such values of pH at which the dissociation is strongest, the chain will tend to elongate because of the electrostatic repulsion between neighboring ionized molecules. If the division of chromosomes affects the pH in such a way as to reduce the dissociation, the chain will coil up under the influence of the thermal agitation which now will prevail over the electrostatic repulsion.

A tendency toward parallel orientation may also be produced by van der Waalsian attraction forces between the end molecules of each chain.

Let us now see what quantitative conclusions could be drawn from the above picture.

The reader who is familiar with the work on the kinetics of long molecules will notice that the following discussion can readily be improved and refined by the use of methods developed by other authors. In particular, the problem of "relaxation time" can be treated rigorously by the method used by J. G. Kirkwood and R. Fuoss (1941). However at the present stage of our development we are interested only in very rough approximations. There are practically no *quantitative* data on mitosis with which to compare any exact mathematical conclusions. Therefore we shall confine ourselves to much simpler methods which give us only orders of magnitudes. The outlined treatment will have to be mathematically refined in future developments.

Consider a long molecular chain consisting of N links, each of length l. In the absence of any external or internal fields, the molecule, as shown by W. Kuhn (1934) will have a "knotted" shape due to thermal agitation. The probability $W(r_1)dr_1$ that the ends A and A' of the molecular chain are at a distance between r_1 and $r_1 + dr_1$, is given by

$$W_1(r_1)\,dr_1 = \left(\frac{1}{b\sqrt{2\pi N}}\right)^3 e^{-\frac{r_1^2}{2b^2N}}\, 4\pi r_1^2 dr_1, \tag{1}$$

where

$$b = \frac{1}{\sqrt{3}}\sqrt{\frac{1+\cos\beta}{1-\cos\beta}}\, l, \tag{2}$$

β denoting the valence angle between two consecutive links. Thus b is of the same order of magnitude as l.

The probability that the middle point of the chain is at a distance between r_2 and $r_2 + dr_2$ from the line AA' is

$$W_2(r_2)\,dr_2 = \frac{1}{b^2}\frac{4}{N}\, e^{-\frac{2r_2^2}{b^2N}}\, r_2 dr_2. \tag{3}$$

Finally, Kuhn introduces a "third dimension" of the molecule in the direction normal to the r_1, r_2 plane. The probability in this direction is given by

$$W_3(r_3)\,dr_3 = \frac{4}{b\sqrt{N\pi}}\, e^{-\frac{4r_3^2}{b^2N}}\, dr_3. \tag{4}$$

The average values of r_1, r_2 and r_3, and the mean squares are given by

$$\bar{r}_1 = 2\sqrt{\frac{2}{\pi}N\,b}\,; \qquad \bar{r}_2 = \tfrac{1}{2}\sqrt{\frac{\pi}{2}N\,b}\,; \qquad \bar{r}_3 = \frac{b}{2\sqrt{\pi}}\sqrt{N}\,; \tag{5}$$

$$\overline{r_1^2} = 3Nb^2; \qquad \overline{r_2^2} = \tfrac{1}{2}Nb^2; \qquad \overline{r_3^2} = \frac{b^2}{8}N. \tag{6}$$

If $l \sim 10^{-8}$ cm, then, in order to have a chain of a length $\sim 10^{-3}$ cm, which is the size of the achromatic spindle, we must have $N \sim 10^5$. For such long chains we may apply with Kuhn the concept of entropy to a single chain. Denoting by k the Bolzmann constant, we have for the entropy of a chain

$$s = s_1 + s_2 + s_3 = k[\log W_1(r_1) + \log W_2(r_2) + \log W_3(r_3)]. \tag{7}$$

Hence denoting by c_1, c_2 and c_3 three constants we have

$$s_1 = c_1 + 2k\log r_1 - \frac{kr_1^2}{2b^2N}; \tag{8}$$

$$s_2 = c_2 + k\log r_2 - \frac{2kr_2^2}{b^2N}; \tag{9}$$

$$s_3 = c_3 - \frac{4kr_3^2}{b^2N}. \tag{10}$$

The total work done in stretching a chain from its average length $\bar{r}_1$ to its full length lN is given by the corresponding change Δf of the free energy. Since

$$f = u - Ts\,, \tag{11}$$

where u is the total energy and T is the absolute temperature, and since we consider now the case in which $\Delta u = 0$, therefore

$$\Delta f = -T\Delta s = -T(\Delta s_1 + \Delta s_2 + \Delta s_3). \tag{12}$$

With $b \sim 10^{-8}$ cm, $N \sim 10^5$, we find that $\bar{r}_1 \sim 5 \times 10^{-6}$, $\bar{r}_1^2 \sim 3 \times 10^{-11}$ and, from equation (8), $\Delta s_1 \sim -10^{-11}$ erg degree^{-1}. The contribution of the log r_1 term is insignificant. As the molecule is stretched completely the "cross-dimensions" $\bar{r}_2$ and $\bar{r}_3$ change from their initial values to about 10^{-8} cm. With the same values as before we find that $\Delta s_2 \sim \Delta s_3 \sim 10^{-16}$ erg degree^{-1}. Hence in computing the work and the forces, we may neglect the terms s_2 and s_3 in equation (7) and put approximately

$$s = s_1\,. \tag{13}$$

From equation (12) it follows that the force of contraction of a stretched chain is given by $F = -T(\partial s/\partial r_1)$ and because of (13) and (8) we find

$$F = -kT\left(\frac{2}{r_1} - \frac{r_1}{b^2N}\right). \tag{14}$$

For sufficiently large extensions ($r_1 > 10^{-5}$ cm) the first term becomes negligible and the force varies approximately linearly with the extension.

If such a stretched chain is left to itself it will again return, due to thermal agitation, to the "knotted" configuration described by equations (1), (3), and (4). The rate of return may be approximately evaluated in the following way: Consider each link of the chain as a sphere of radius $a = 10^{-8}$ cm. While the ends of the chain move by a given amount toward each other, the middle link remains approximately undisplaced, while the velocity of the other links *with respect to the medium* varies between $(1/2)(dr_1/dt)$ and zero. Hence, approximately, we may put the average velocity of each link with respect to the medium as equal to $v = -(1/4)(dr_1/dt)$. Assuming Stokes' law, the average force of resistance on each link equals $6\pi\eta av$ where η is the viscosity of the medium. The total force of resistance to the N links is then $6\pi\eta avN$, or $-\dfrac{3}{2}\pi\eta aN(dr_1/dt)$. Neglecting inertial terms, this must be equal to the force F as given by equation (14) in which, as we have seen, we may neglect the term $2/r_1$. We thus find

$$\frac{3}{2}\pi\eta aN\frac{dr_1}{dt} = -\frac{kT}{b^2N}r_1. \tag{15}$$

Integrating and taking $r_1 = lN$ for $t = 0$ (completely stretched chain) we find

$$r_1 = lNe^{-(2kT/3\pi\eta ab^2N^2)t}\,. \tag{16}$$

For values of a, b and N assumed above, for $T \sim 300°$, and $\eta \sim 1$ abs. c.g.s units, the value of the exponent is about 1 sec^{-1}. Thus after one second the chain will have "coiled up" to only $1/e$ of its original length and the "relaxation time" is thus of the order of seconds.

For a constant length of the chain $b^2N^2 =$ const. Hence the relaxation time increases with $a \sim b$. For $a \sim 10^{-6}$ cm, $N \sim 10^3$, and $\eta \sim 1$ c.g.s units, we find for the relaxation time a value of about

100 sec. With the same values of a and N, but with $\eta \sim 10^{-2}$, it again reduces to about 1 second.

The total work of stretching a chain is, according to equation (14), approximately equal to $kTl^2N^2/2b^2N \sim kTN/2$. The energies of molecular interactions which may depend on their relative orientations are of the order of $kT \sim 10^{-13}$ erg, frequently several times that amount. Hence if each link contains a molecule or group of molecules of a polar nature which tends to orient the corresponding link of another chain in a parallel direction, the total decrease of the potential energy when two chains are stretched and arranged side by side parallel to each other may well be sufficient to supply the work $kTN/2$ necessary for stretching each chain. In that case a spontaneous orientation would occur. However, the intermolecular forces which tend to produce any orientation are of short range and become negligible at distances exceeding 10^{-7} cm. Hence if we had only f-fibers, which would number about 10—100 for the whole nucleus, no orienting effect would occur because the average distance between such fibers would be about 10^{-4} cm, which is much too far away for any ordinary intermolecular forces to become effective. As we said above, however, we may have a large number of f_1-molecules. If their number equals or exceeds 10^{10}, the average distance will be of the order of 10^{-8} cm, and an orientation may occur.

The orienting forces may act like a "molecular field," similar to the one introduced by P. Weiss (1907) in the theory of ferromagnetism. The more each chain is stretched, the stronger may be the orienting force, and that in its turn will tend to stretch the chains further. Thus, approximately, we shall have a "stretching force" F_s acting on each molecule and proportional to r_1 (which we may consider as approximately the same for all molecules). Thus

$$F_s = -Ar_1 , \tag{17}$$

the minus sign indicating that the force is a stretching one. For a spontaneous orientation to occur we must have $F_s > F$, or

$$A - (kT/b^2N) > 0 . \tag{18}$$

In that case we may roughly estimate the rate of orientation. The viscous resistance is again equal to $(3/2)\pi\,\eta\,aN(dr_1/dt)$. This must be equal to $[A - (kT/b^2N)]\,r_1$. Thus we find now

$$\frac{3}{2}\pi\,\eta\,aN\frac{dr_1}{dt} = \left(A - \frac{kT}{b^2N}\right)r_1 . \tag{19}$$

Putting now for $t = 0$, $r_1 = \bar{r}_1$, as given by equation (5), and introducing the notation

$$B = \frac{2}{3} \frac{A - (kT/b^2N)}{\pi \eta a N}, \tag{20}$$

we find

$$r = \bar{r}_1 e^{Bt}. \tag{21}$$

Because of (18), $B > 0$. The absolute value of B may, however, be much smaller than kT/b^2N. In that case we will find much smaller values for B than those found for the exponent in equation (16). The time for complete orientation may be as high as many minutes or even hours.

Thus with plausible values for the physical constants involved, we see that the computed time of formation of the mitotic spindle and of its disappearance is of the correct order of magnitude.

If upon division of the chromosomes the substance S is produced, which reduces the orienting intermolecular forces, all the fibers will contract and the daughter chromosomes will be pulled apart. A number of interesting possibilities may be considered, depending on whether the f–fibers are first "loosened" from the f_1– matrix and contract alone, or whether the f_1–fibers also begin to contract. Considering a chromosome, as a very rough approximation as a sphere of radius $a_1 = 10^{-4}$ cm, remembering that it is pulled by an f–fiber which has only $N/2$ links, and considering that the other end of the fiber is rigidly attached to the end of the spindle, we find, instead of equation (15)

$$(\tfrac{3}{4} \pi \eta a N + 6 \pi \eta a_1) \frac{dr_1}{dt} = -\frac{2kT}{b^2N} r_1. \tag{22}$$

With the same values of the constants as before this gives us again roughly one second for the movement of the daughter chromosomes from equator to the pole. This is too short a time. It will, however, be considerably lengthened if we consider that the S–substance is not likely to abolish the orienting forces at once, but rather weaken them gradually.

One of the quantitative consequences of this theory, which can be checked by proper experiments, is the variation of the distance of the daughter chromosome from the pole, as given by equation (22). Data for that are, however, not available now.

LITERATURE

Kirkwood, J. G. and Fuoss, R. M. 1941. "Anomalous Dispersion and Dielectric Loss in Polar Polimers." *Jour. Chem. Physics*, 9, 329–340.

Kuhn, W. 1934. "über die Gestalt fadenförmiger Moleküle." *Koll. ZS.*, 68, 2–15.

Kuhn, W. 1936. "Beziehungen zwischen Molekülargrösse, statistischer Molekülargestalt und elastischen Eigenschaften hochpolymerer Stoffe." *Koll. ZS.*, 76, 258–271.

Landahl, H. D. 1942. "A Mathematical Analysis of Elongation and Constriction in Cell Division." *Bull. Math. Biophysics*, 4, 45–62.

Rashevsky, N. 1941. "Some Remarks on the Movement of Chromosomes During Cell Division." *Bull. Math. Biophysics*, 3, 1–3.

Rashevsky, N. 1938. *Mathematical Biophysics.* Chicago: University of Chicago Press.

Rashevsky, N. 1948. *Mathematical Biophysics.* Revised Edition. Chicago: University of Chicago Press.

Schrader, F. 1944. *Mitosis: The Movement of Chromosomes in Cell Division.* New York: Columbia University Press.

Weiss, Pierre. 1907. "L'Hypothèse du Champ Moléculaire et la Propriété Ferromagnétique." *Jour. de Physique*, (4), 6, 661–690.

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 11, 1949

OUTLINE OF A PROBABILISTIC APPROACH TO ANIMAL SOCIOLOGY: I

ANATOL RAPOPORT
COMMITTEE ON MATHEMATICAL BIOLOGY
THE UNIVERSITY OF CHICAGO

The non-transitive character of the peck right relation gives rise to different "social structures" in an aggregate of individuals. A method is developed for computing the respective probabilities of occurrence for each type of structure in small aggregates on the basis of random determination of peck right between each pair of individuals.

Whether sociology can ever become a full-fledged "science" (a description of a class of events predictable on the basis of deductions from a constant *rationale*) depends on whether the terms which sociologists employ to describe events can be analyzed into quantifiable observables.

As long ago as the XVII century attempts were made to construct a "social physics" (Murchison, 1935). These attempts, inspired, no doubt, by the magnificent successes of applying the quantitative methods to "natural" (that is, simpler) phenomena, failed, because, as Murchison so pithily put it, the "social physicists" made the "mistake of trying to do *what* physics has done, instead of trying to do *as* physics has done." The major task of the modern mathematical sociologist (or the mathematical biologist, for that matter) is to rectify that mistake, that is, to find a way of applying the methods of physical science to social phenomena without trying to force them into the known frameworks of physics by naive analogies.

If one asks *what* physics has done, one might answer that it has discovered the inverse square laws of gravitation and electrostatics, the principles of conservation of mass and energy and their generalization, invented descriptions of subatomic events in terms of wave and quantum mechanics, etc. But the answer to the question, *how* physics has done, is altogether different. Physics has begun by observation and by more or less quantitative descriptions of the results of observation. Then it has sought to establish systems of *covariables,* in other words, to reduce the description of events to "if so, then so" form. Finally, it established a *rationale,* that is, a mathematical formulation of "laws" from which all past and many future observations could be deduced. To do as physics has done means to pursue this program. As Murchison says,

> The first major task is to identify simple quanta of social phenomena that by definition are describable in terms of time and space categories. After the identification comes the second major task of developing a method of measuring these quanta in terms of such categories. The third major task involves the application of the technique of covariables. The fourth major task has to do with the formulation of law and theory.

Recognizing that the social behavior of non-humans is almost certainly simpler than that of humans, Murchison proceeded to identify "quanta of social phenomena" in the *Gallus domesticus* (the common chicken), which, as he points out, is a laboratory animal with a long and distinguished history. The "quanta" identified by Murchison were the so-called Social Reflex No. 1, observed in the movement toward another member of the same species, and Social Reflex No. 2, fighting a member of the same species. He then established methods for measuring these quanta (Reflex No. 1 by the velocity of approach and Reflex No. 2 by willingness to fight and the success in battle). As covariables Murchison considered mass, momentum per second and kinetic energy per second in Reflex No. 1, and their relation to success in battle. Another covariable considered was "social rank," that is, the peck order position of the bird in its own flock.

N. Collias (1942) points out that among hens "social rank" is established by the result of the first encounter. He then proceeds to analyze the correlations between success in combat with a stranger and other measurable characteristics, such as size of comb, degree of moulting, mass, and social rank in the home flock. Thus the major tasks of identifying, measuring, and correlating "social quanta" are under way in the study of events, which may well be considered simple social phenomena. In these studies, the unit is taken to be the *individual,* and the observables are his individual characteristics and his relations to other individuals within and outside his society (flock).

We wish to attack this problem from a somewhat different point of view and by a different method. Our unit will be the *society* (a small flock of birds) and the observables will be the elements of structure of that society. Our method will be the method of theoretical science: "mental experimentation," applied to the search for a rationale. As has been repeatedly pointed out by N. Rashevsky (1947) and others, one need not insist too rigidly on "getting sufficient facts" before attempting to formulate theories, if the "theoretical" nature of theories is constantly borne in mind. In fact "mental experimentation" can often point the way to a fruitful experimental program.

The search for a rationale is in a way analogous to an experimental program, except that the experiments are performed with paper and pencil instead of with subjects, materials, and apparatus. Even rationales which give "negative results," i.e., fail to account for the observed relations, have a certain degree of scientific value. Often by observing the "degree of unworkability" of a hypothesis, indications are obtained on how to proceed in a search of a better rationale.

In the work of Murchison, Collias, and others, where correlation is sought between observables, it is desirable to reduce unknown and variable factors to a minimum, since these tend to reduce the significance of correlations. But there is another approach, the probabilistic, where if the variables are unknown enough, numerous enough, a certain randomness may be assumed, and the problem may be formulated in terms of random events. We shall adopt this approach. Since the principles developed here are applicable to any analogous situation, we shall speak of societies of "individuals" instead of chickens.

General Considerations.

Peck order or "peck right" is a binary asymmetric relation between each pair in a finite aggregate of individuals (a society). The relation will be designated by $>$ or by $<$, so that $A > B$ is read "A pecks B" and, of course, $A < B$ is read "A is pecked by B." The relation is asymmetric—if $A > B$ holds, then $B > A$ does not hold, at least not simultaneously. Two facts about this relation (as observed in bird and other societies) are of interest to the mathematical biologist.

1. The relation is not necessarily transitive—$(A > B) \cdot (B > C)$ does not always imply $A > C$.
2. The set of relations in a given society may change with time, tending to transform the structure of the society into a simple chain (Allee, 1931; Murchison, (1935) thus,

$$A > B > C > \cdots > Z. \tag{1}$$

The first observation points to indeterminate factors in the dynamics of such a society, while the second points to determinate factors. It is, of course, possible to postulate a deterministic dynamics and still get intransitive relations. If peck right is completely determined by relations existing between certain inherent or acquired characteristics of the individuals, one may formulate the determination of peck right thus

$$\begin{aligned} A > B &\rightleftarrows f(A\,,B) > f(B\,,A)\,; \\ B > A &\rightleftarrows f(B\,,A) > f(A\,,B), \end{aligned} \tag{2}$$

where f is a certain function, whose arguments are the respective "characteristics" of the individuals involved in the peck right relation, and arrows indicate implication.

Now if f were a function of a single variable, then the implication

$$[f(X) > f(Y)] \cdot [f(Y) > f(Z)] \rightarrow f(X) > f(Z) \tag{3}$$

would have to hold, and peck right would have to be always transitive. If, however, f is a function of more than one variable, then it is easy enough to construct functions f, such that

$$\begin{gathered} f(A\,,B) > f(B\,,A)\,; \qquad f(B\,,C) > f(C\,,B)\,; \\ f(C\,,A) > f(A\,,C). \end{gathered} \tag{4}$$

One need only to mark the points $(A\,,\,B)$, $(B\,,\,A)$, $(B\,,\,C)$, $(C\,,B)$, $(A\,,\,C)$, and $(C\,,\,A)$ in a plane, construct perpendiculars at these six points of lengths $f(x\,,\,y)$ such that the inequalities (4) hold, and pass an arbitrary surface through the six points $f(x\;,\;y)$. If the equation of that surface is given by $z = f(x\,;\,y)$, and if peck right between X and Y is determined by whether $f(X\,,\,Y) \gtrless f(Y\,,X)$, one can see at once that for the individuals A, B, and C the peck right will not be transitive.

One could, therefore, begin the logical development of the peck right problem by examining some such function $f(x\,,\,y)$ which would allow both transitive and intransitive relations, depending on the values of its arguments. The structure of a society will then be determined by the form of f and by the distribution of the "characteristics" among its members. Any change in the structure will have to be derived either by taking the characteristics to be functions of time or by assuming changes in the form of f with time.

This approach does not seem attractive, since the number of functions one may choose for f is prodigious, and one has no *a priori* reason to prefer one to another.

The probabilistic hypothesis, on the other hand, supposes that the structure of a society is determined *to a certain extent* by the outcomes of chance events. This means, in the last analysis, that we are taking into account our ignorance of the determining factors in the dynamics of the society, but we are formulating this ignorance mathematically as is done in any probabilistic approach to a problem. It is also possible to combine deterministic with probabilistic factors, where, for example, one assigns to events probabilities which are functions of some known quantities.

There is another reason besides mathematical convenience for introducing chance events into the dynamics of animal sociology. If, as is indicated by the observation 2 above, the later stages of a society are more "organized" than the early stages, one could postulate the initial working of chance and an inherent bias in the situation which causes the structure to "gravitate" to a certain form. Thus, if one spills peas on an uneven surface, the initial distribution of the peas will be nearly random, but eventually they will assume positions along the lines of minimum potential energy.

A natural choice of a chance event is the result of an encounter between two individuals. We shall assume that in any such encounter, one will enjoy "victory" and the other suffer "defeat." It will be further assumed that the probabilities of the outcomes of such encounters can be computed on the basis of the knowledge of the history of the individuals and their inherent characteristics. The "structure" of a society, that is, the distribution of social rank will be supposed to be determined by the outcomes of such encounters.

One can make several different hypotheses about factors influencing the outcomes and about the factors resulting from them. The following are examples of such hypotheses, not necessarily consistent with each other.

1. The result of the first encounter between any two individuals is equiprobable (probability for victory $= 1/2$ for each).
2. The probabilities of the results of the first encounter are functions of the respective characteristics (independent of time) of the individuals concerned.
3. The probability of the result of each encounter depends on the result of the immediately preceding encounter between the same individuals.
4. The result of each encounter depends on the total history of encounters experienced by the individuals concerned.
5. Relative social rank between two individuals is completely determined by the result of their first encounter.
6. The result of each encounter determines the relative social rank between two individuals, but other encounters may ensue in which the relative social rank may be reversed. The probability of

the reversal depends on the difference in the social rank between the individuals.

8. The probability of an encounter taking place at all is a function of the relative social rank.

Etc.

One could begin with any self-consistent combination of these hypotheses and derive its implications. These implications would be conclusions about the probability distributions of various "types" of societies, and, if the frequency of encounters were also known or assumed, one could (in principle) derive expressions for the changes of these probability distributions with time. One should thus obtain a "wave" of probability distribution. The abscissa of this wave would be a particular type of structure of a society (to be defined below); the ordinate, the probability frequency of this structure; and the whole wave would be moving into the third (time) dimension, changing its shape as it moved, so that after a long time it would exhibit a sharp crest at the type of structure toward which the society gravitates.

Thus a surface would be generated by this probability wave which could be considered as a complete representation of the "statistical dynamics" of the society.

It remains to define the "type of structure" or simply the "structure" of a society. Evidently a definition is desired which would enable us to recognize structure by observation. In a society of N individuals, each will have $N - 1$ peck right relations of which r will be dominant (the right to peck) and $N - 1 - r$ submissive (the necessity to be pecked). There will then be a distribution of numbers $(r_1, r_2, \cdots r_N)$ among the N individuals such that

$$\sum_{i=1}^{N} r_i = \tfrac{1}{2}N(N-1). \qquad (5)$$

This set of numbers $(r_1, r_2, \cdots r_N)$ together with all its permutations will be defined as a *structure* of the society. A structure can also be represented by a diagram where each individual is placed on a level determined by the number r_i of dominant peck right relations he enjoys. Thus a society of four individuals can have exactly four structures, diagramatically represented in Figure 1. Arrows indicate peck right; the numbers indicate the dominant peck right relations of each individual. The problem of determining the dynamics of such a society, given a set of assumptions about the occurrence and results of encounters, reduces to the calculation of the probability of occurrence

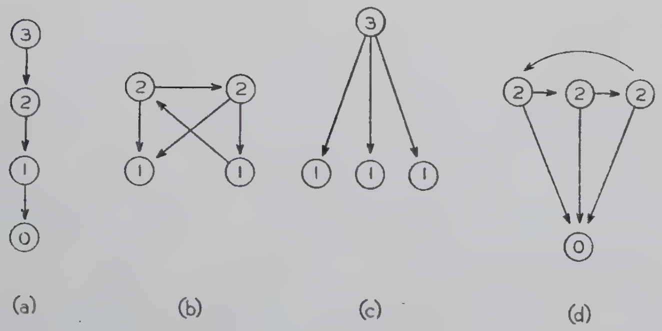


FIGURE 1

of each structure at each time in the history of the society. In this paper the following simple assumptions will be made.

1. The results of the first encounter between any two individuals are equiprobable.
2. The result of the first encounter determines permanently the peck right between the individuals concerned (i.e., the winner will always peck the loser).

Thus the shape of the probability distribution wave is permanently established (independent of time) as soon as all the $\frac{1}{2}N(N-1)$ encounters have taken place. An alternative interpretation is that we are limiting our observations to the period when the structure depends on the results of first encounters only.

Three Individuals.

Since only one type of structure is possible for a society of two individuals, namely, $(1, 0)$, we shall begin with the case of three individuals.

There are $2^3 = 8$ possible sets of outcomes of the three encounters. Each outcome leads to one of two possible structures, $(2, 1, 0)$ and $(1, 1, 1)$. The respective probabilities of these structures are 3/4 and 1/4, since six of the eight outcomes map on the first (simple chain), and two on the second (simple cycle). This solves the problem of the three individuals.

Some Aspects of the N-*Individual Problem.*

One would like to generalize the solution to N individuals. However, difficulties are encountered even under the extremely simple assumptions made above.

The probability of an event is defined as the ratio of the number of ways the event can happen as a result of certain occurrences to the total number of possible occurrences, assumed equally likely. Now in a society of N individuals there will eventually take place $\frac{1}{2}N(N-1)$ first encounters. Since there are two possible results to each encounter, the number of all possible occurrences (where an occurrence is the complete set of results) will be $2^{\frac{1}{2}N(N-1)}$. To compute the probability of a certain structure, one must calculate how many of those "sets of results" will "map" on a particular distribution of peck rights $(r_1, r_2, \cdots r_N)$ or on any permutation thereof. This would be a simple matter if there were a one-to-one correspondence between each set of results and each *ordered* set $(r_1, r_2, \cdots r_N)$. Then the probability of the structure $(r_1, r_2, \cdots r_N)$ would be simply the number of *distinguishable* permutations of the set divided by $2^{\frac{1}{2}N(N-1)}$. That the mapping is not in general one-to-one is shown by the following counterexample.

Each set of results can be represented by a skew-symmetric matrix, for example, in a society of four individuals by

$$\begin{array}{c|cccc|} & A & B & C & D \\ A & 0 & -1 & -1 & +1 \\ B & +1 & 0 & +1 & -1 \\ C & +1 & -1 & 0 & -1 \\ D & -1 & +1 & +1 & 0 \end{array} \qquad (6)$$

where the article a $+1$ in the X-row, Y-column indicates $X > Y$ and a -1, $X < Y$. Note that the matrix (6) maps on the structure $(1, 2, 1, 2)$. But so does the matrix

$$\begin{array}{c|cccc|} & A & B & C & D \\ A & 0 & -1 & +1 & -1 \\ B & +1 & 0 & -1 & +1 \\ C & -1 & +1 & 0 & -1 \\ D & +1 & -1 & +1 & 0 \end{array} \qquad (7)$$

Without solving the problem of determining the probability distribution of structures for all N, it is nevertheless possible to reword it in such a way as to state the results in terms of properties of such skew-symmetric matrices. In this form the problem may, perhaps, prove susceptible to attack by methods derived from the theory of these matrices. For this reason, we introduce an alternative definition of structure. Let the structure of a society be defined by a set of numbers

$$(r_1', r_2' \cdots r_N'), \qquad (8)$$

each of which represents the excess of the number of dominant relations over the number of submissive relations associated with an individual. Then evidently

$$\sum_{i=1}^{N} r_i' = 0. \tag{9}$$

Theorem I. *In a society of* N *individuals the structures defined by* $(\mathrm{r}_1, \mathrm{r}_2, \cdots\cdot \mathrm{r}_N)$ *and* $(\mathrm{r}_1', \mathrm{r}_2', \cdots\cdot \mathrm{r}_N')$ *are in one-to-one correspondence.*

Proof. Let a given structure be given by $(r_1, r_2, \cdots\cdot r_N)$ in the sense of the first definition and by $(r_1', r_2', \cdots\cdot r_N')$ in the sense of of the second. Then

$$r_i' = r_i - (N - 1 - r_i) = 2r_i - N + 1. \tag{10}$$

Also

$$r_i = 1/2 (r_i' + N - 1). \tag{11}$$

The right-hand side of equation (11) always represents an integer since from (10) we see that r_i' is even or odd depending on whether N is odd or even. Thus the r_1' are uniquely determined by the r_i and vice versa. This proves the theorem.

Consider now the class $\mathcal{M}$ of all skew-symmetric matrices (a_{ij}) of order N, all of whose non-diagonal elements are units, that is, $a_{ij} = 1$ or -1; $a_{ij} = -a_{ji}$.

Theorem 2. *Let* $(\mathrm{a}_{\mathrm{ij}})$ *be a matrix of the class* $\mathcal{M}$, *and let*

$$\mathrm{S}_\mathrm{k} = \left(\sum_{\mathrm{j}=1}^{\mathrm{N}} \mathrm{a}_{1\mathrm{j}}, \sum_{\mathrm{j}=1}^{\mathrm{N}} \mathrm{a}_{2\mathrm{j}}, \cdots\cdot \sum_{\mathrm{j}=1}^{\mathrm{N}} \mathrm{a}_{\mathrm{Nj}} \right)$$

be the ordered set of its row sums. Let p_k *be the number of distinguishable permutations of the set* S_k, *and* m_k *the number of matrices in* $\mathcal{M}$ *giving rise to the set of row sums* S_k. *Then the probability of the structure* $(\mathrm{r}_1', \mathrm{r}_2', \cdots\cdot \mathrm{r}_N')$, *where* $\mathrm{r}_\mathrm{i}' = \sum_{\mathrm{j}=1}^{\mathrm{N}} \mathrm{a}_{\mathrm{ij}}$ *is given by*

$$P(r_1', r_2', \cdots\cdot r_N) = 2^{-\frac{1}{2}N(N-1)}\, p_k m_k. \tag{12}$$

Proof. Each matrix (a_{ij}) represents an equiprobable set of encounter outcomes. Furthermore, the row sums of the matrix (a_{ij}) represent the excess of victories over defeats of each individual and therefore, according to our assumptions, the excess of his dominant over his submissive relations. Therefore m_k represents the number of sets of outcomes mapping on a particular ordered set S_k. But any permutation of S_k, $S_k^{(i)}$ can be interpreted as simply a re-labeling of the individuals. Hence the number of matrices mapping on $S_k^{(i)}$ $(i = 1, \cdots p_k)$ is likewise m_k. Moreover sets of outcomes mapping on distinct sets S are certainly distinct. The total number of matrices in $\mathscr{M}$ for a given N is $2^{\frac{1}{2}N(N-1)}$. The number of distinct matrices mapping on the *unordered* set S_k is $p_k m_k$. This proves the theorem.

Because of the one-to-one correspondence between the structures $(r_1, r_2, \cdots r_N)$ and $(r_1', r_2', \cdots r_N')$ established in Theorem 1, we have the

Corollary. *The probability of a structure* $(\mathrm{r}_1, \mathrm{r}_2, \cdots \mathrm{r}_N)$ *is also* $2^{\frac{1}{2}\mathrm{N}(\mathrm{N}-1)}\, \mathrm{p}_k \mathrm{m}_k$.

A general method of finding m_k for any S_k and any N is at this time unknown to the author. However, for N not too large, say $N \leqq 6$, the probabilities of all structures can be easily computed by a systematized procedure, which will now be described.

First let us consider a procedure for writing down all possible structures $(r_1, r_2, \cdots r_N)$ of a society of N individuals. The simplest such structure is a simple chain (complete hierarchy). It is represented by the set of numbers

$$(N-1, N-2, \ldots\ldots. 0). \qquad (13)$$

This structure is a mapping of the matrix

$$\left\| \begin{matrix} 0 & + & + & \ldots. & + \\ - & 0 & + & \ldots. & + \\ - & - & 0 & \ldots. & + \\ \vdots & & & \ddots & \\ - & - & - & \ldots. & 0 \end{matrix} \right\| \qquad (14)$$

where all the pluses are above the diagonal (the units have been replaced by their signs for simplicity). All other matrices of class $\mathscr{M}$ are obtained by changing a set of pluses in matrix (14) to minuses and the symmetrically situated minuses to pluses, i.e., by a set of interchanges of pluses and minuses. But not all such interchanges will give rise to structures different from (13), because some of them may map on permutations of (13), which, by definition, represent the same structure. Such permutations can be avoided if we consider only such interchanges which preserve the order of magnitudes in the $\sum_{n=1}^{N} a_{ij} = r_i$, that is, their non-increasing order.

Such interchanges can be represented as follows. Consider N boxes containing respectively $N-1, N-2, \cdots 0$ balls. We are interested in the operations which throw the set $N-1, N-2, \cdots 0$ into another set $(r_1, r_2, \cdots r_N)$, where $r_i \geqq r_j$ for $i < j$, and such that $\sum_{i=1}^{N} r_i = \frac{1}{2}N(N-1)$. Such operations can be accomplished by transferring the balls in several ways, but if the balls are not distinguishable, for each transformation

$$(N-1, N-2, \cdots 0) \rightarrow (r_1, r_2, \cdots r_N) \tag{15}$$

there exists a unique, *most direct* way of transferring the balls, where each box either gives up balls or receives them or neither, but not both. For instance the transformation $(3, 2, 1, 0) \rightarrow (2, 2, 1, 1)$ can be accomplished by transferring a ball from the first box to the second, from the second to the third, and from the third to the fourth. But there is a unique *direct* transfer with the same result, namely where a ball is taken from the first box and put into the fourth. It is also evident that all such direct order-preserving transfers are from left to right. But each transfer represents an interchange of signs in matrix (14) and also avoids the permutations of (r_i). Hence we have

Theorem 3. *The number of possible structures of a society of* N *individuals equals the number of ways direct transfers of balls can be made preserving non-increasing order in a set of* N *boxes containing* N — j *balls respectively* (j = 1 ···· N).

By systematizing these transfers in some convenient manner, we can write down all possible structures of societies of N individuals, if N is not too large. Thus a society of six individuals gives rise to 20 possible structures. Beginning with the complete hierarchy (5, 4, 3, 2, 1, 0) the other structures can be written down by imagining all transfers of one ball, then all transfers of two balls, etc., as shown in Table 1.

TABLE 1

S_1: (5, 4, 3, 2, 1, 0)
S_2: (4, 4, 4, 2, 1, 0)
S_3: (4, 4, 3, 3, 1, 0)
S_4: (4, 4, 3, 2, 2, 0)
S_5: (4, 4, 3, 2, 1, 1)
S_6: (5, 3, 3, 3, 1, 0)
S_7: (5, 3, 3, 2, 2, 0)
S_8: (5, 3, 3, 2, 1, 1)
S_9: (5, 4, 2, 2, 1, 1)
S_{10}: (5, 4, 2, 2, 1, 1)
S_{11}: (5, 4, 3, 1, 1, 1)
S_{12}: (4, 3, 3, 3, 2, 0)
S_{13}: (4, 3, 3, 3, 1, 1)
S_{14}: (4, 3, 3, 2, 2, 1)
S_{15}: (4, 4, 2, 2, 2, 1)
S_{16}: (5, 3, 2, 2, 2, 1)
S_{17}: (3, 3, 3, 3, 2, 1)
S_{18}: (3, 3, 3, 2, 2, 2)
S_{19}: (5, 2, 2, 2, 2, 2)
S_{20}: (4, 3, 2, 2, 2, 2)

It remains to indicate a method for computing the probabilities of structures. We have seen that these probabilities are determined by the number of matrices of class $\mathcal{M}$ which map on the set $(r_1, r_2, \cdots r_N)$ or any of its permutations. Really it is necessary only to determine the number of matrices mapping on the *ordered* set $(r_1, r_2, \cdots r_N)$, as indicated in the proof of Theorem 2. The total number of matrices mapping on a structure will then be the number mapping on an *ordered* set $(r_1, r_2, \cdots r_N)$ times the number of distinguishable permutations of this set.

The society of three individuals has already been described. For the case of four individuals, we have $2^6 = 64$ possible matrices. The ordered set $(3, 2, 1, 0)$ can be obtained only from the matrix

$$\begin{Vmatrix} 0 & + & + & + \\ - & 0 & + & + \\ - & - & 0 & + \\ - & - & - & 0 \end{Vmatrix}$$

This set has 24 distinguishable permutations. Hence the probability of the complete hierarchy in a society of four individuals is

$$P(3, 2, 1, 0) = 24/64 = 3/8 .$$

The ordered set $(2, 2, 2, 0)$ can be obtained in two different ways,

$$\begin{Vmatrix} 0 & + & - & + \\ - & 0 & + & + \\ + & - & 0 & + \\ - & - & - & 0 \end{Vmatrix} \quad \text{and} \quad \begin{Vmatrix} 0 & - & + & + \\ + & 0 & - & + \\ - & + & 0 & + \\ - & - & - & 0 \end{Vmatrix}$$

Its distinguishable permutations are four. Therefore

$$P(2, 2, 2, 0) = 2 \times 4/64 = 1/8 .$$

The set $(2, 2, 1, 1)$ can be obtained in four different ways

$$\begin{Vmatrix} 0 & + & + & - \\ - & 0 & + & + \\ - & - & 0 & + \\ + & - & - & 0 \end{Vmatrix} \qquad \begin{Vmatrix} 0 & + & - & + \\ - & 0 & + & + \\ + & - & 0 & - \\ - & - & + & 0 \end{Vmatrix}$$

$$\begin{Vmatrix} 0 & - & + & + \\ + & 0 & + & - \\ - & - & 0 & + \\ - & + & - & 0 \end{Vmatrix} \qquad \begin{Vmatrix} 0 & - & + & + \\ + & 0 & - & + \\ - & + & 0 & - \\ - & - & + & 0 \end{Vmatrix}$$

its distinguishable permutations are six. Therefore

$$P(2,2,1,1) = 4 \times 6/64 = 3/8.$$

Finally (3, 1, 1, 1) can be obtained in two ways

$$\begin{vmatrix} 0 & + & + & + \\ - & 0 & + & - \\ - & - & 0 & + \\ - & + & - & 0 \end{vmatrix} \qquad \begin{vmatrix} 0 & + & + & + \\ - & 0 & - & + \\ - & + & 0 & - \\ - & - & + & 0 \end{vmatrix}$$

with four distinguishable permutations. Hence

$$P(3,1,1,1) = 2 \times 4/64 = 1/8.$$

This completes the case of four individuals. The probabilities of the structures occurring in societies of five individuals can be easily computed by this method. The results are given in Table 2.

TABLE 2

Structure	Matrices Mapping on Ordered Set	Number of Permutations	Probability
(4, 3, 2, 1, 0)	1	120	$2^{-10} \times 120$
(3, 3, 3, 1, 0)	2	20	$2^{-10} \times 40$
(3, 3, 2, 2, 0)	4	30	$2^{-10} \times 120$
(3, 3, 2, 1, 1)	8	30	$2^{-10} \times 240$
(4, 2, 2, 2, 0)	2	20	$2^{-10} \times 40$
(4, 2, 2, 1, 1)	4	30	$2^{-10} \times 120$
(4, 3, 1, 1, 1)	2	20	$2^{-10} \times 40$
(3, 2, 2, 2, 1)	14	20	$2^{-10} \times 280$
(2, 2, 2, 2, 2)	24	1	$2^{-10} \times 24$

The foregoing results implicitly predict the distribution of structures in a large number of societies of a given number of individuals, provided the hypotheses used in their derivations are valid. On the other hand there are strong indications that in any actual society like a flock of chickens significant bias factors are at work. We should not therefore expect the distributions in actual flocks to follow the computed distributions given here. However a liaison between the above theoretical considerations and experimental procedure can be accomplished in a variety of ways.

1. The above results can be used as a theoretical stepping stone for complicating the theory by the introduction of biases known to be important.

2. A crude prediction can be made even on the basis of the results obtained, namely that in a situation where peck right is deter-

mined by the result of the first encounter, the probability distribution will depart from that predicted here so as to weigh more heavily those structures which are closer to complete hierarchy. For example on the basis of random results of encounters, the probabiliy of a complete hierarchy in a society of three individuals is 3/4. With the introduction of some bias, we would expect it to exceed that value.

3. Whereas the statistical work of correlating physical characteristics and social rank with success in combat is most significant when the opponents are *not* evenly matched, the results of the present theory can be best realized when they are matched as evenly as possible. Thus experimental work can be extended to another range.

This work was aided by a grant from the Dr. Wallace C. and Clara A. Abbott Memorial Fund of The University of Chicago.

LITERATURE

Allee, W. C. 1931. *Animal Aggregations. A Study in General Sociology.* Chicago: University of Chicago Press.

Collias, N. 1943. "Statistical Analysis of Factors which Make for Success in Initial Encounters between Hens." *The American Naturalist*, 77, 519–538.

Lienau, C. C. 1947. "Quantitative Aspects of Organization." *Human Biology*, 19, 163–216.

Murchison, C. 1935. "The Experimental Measurment of a Social Hierarchy in Gallus Domesticus." *Jour. Gen. Psychol.*, 12, 3–39.

Rashevsky, N. 1948. *Mathematical Theory of Human Relations.* Bloomington: The Principia Press.

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 11, 1949

CONTRIBUTION TO THE LINKAGE THEORY OF AUTOPOLYPLOIDS: II

HILDA GEIRINGER
WHEATON COLLEGE

In the first part of this paper, hereinafter referred to as I, a general segregation distribution was introduced for autopolyploids ($2s$-ploids) with m loci, and r alleles. Random mating, chromosome segregation, distinct generations, and equal segregation distributions for males and females were assumed. In this second part, using this segregation distribution as a basis, a recurrence formula is established which enables us to compute the distribution of gametes for any generation, if this distribution is known for an initial generation. This initial distribution of gametes is derived from the initial distribution of genotypes. The limit behavior of these distributions is completely described in a general limit theorem which contains as particular cases the limit theorems for diploids with m loci, and for $2s$-ploids with one locus (Geiringer, 1944, 1945, 1948).

1. *The General Recurrence Formula.*

The general recurrence laws for polyploids with m loci will now be derived. We write, as before, for the distribution of gametes and for the distribution of zygotes in the nth generation $p^{(n)} = p$, $w^{(n)} = w$, and $p^{(n+1)} = p'$, $w^{(n+1)} = w'$. We consider a $2s$-ploid with m loci and r alleles. We denote the ith allele corresponding to the kth locus by a_{ik} $(i = 1, \cdots\cdot m, k = 1, \cdots\cdot r)$ and the distribution of gametes by

$$p(a_{11}^{z_{11}} \cdots\cdot a_{1r}^{z_{1r}}, \cdots\cdot a_{m1}^{z_{m1}} \cdots\cdot a_{mr}^{z_{mr}}),$$
$$\left(\begin{array}{c} z_{i1} + \cdots\cdot + z_{ir} = s \\ i = 1, \cdots\cdot m \end{array} \right).$$

If $r > s$ some of the z_{ik} will necessarily be zero. There are

$$S = \frac{s!}{z_{11}! \cdots\cdot z_{1r}!} \cdots\cdot \frac{s!}{z_{m1}! \cdots\cdot z_{mr}!}$$

such gametic probabilities for a set of z_{ik} and we want to express $S \cdot p'(a_{11}^{z_{11}} \cdots\cdot a_{1r}^{z_{1r}} \cdots\cdot a_{m1}^{z_{m1}} \cdots\cdot a_{mr}^{z_{mr}})$ in terms of the w' as was done in the case $s = m = 2$ (section 5 of I).

We shall need the marginal distributions of the p-distribution. The definition is analogous to that in the particular case considered in section 5 of I. A general example of marginal summation will suffice. Consider $m=3$ loci, $r=4$ alleles and $s=6$. Denote, in this example, the alleles by $a_1, \cdots, a_4, b_1, \cdots, b_4, c_1, \cdots, c_4$. The p-distribution is a function of eighteen variables each of which takes on four discrete values. From this distribution we may derive marginal distributions of order $17, 16, \cdots, 1$. There are three of order 17, six of order $16, \cdots$, three of order 1. As an example consider the marginal distribution of order 11: $p(a_1^2a_2b_1b_3^3c_1c_2c_4^2)$. We will use summation letters: α taking on the values $a_1, \cdots, a_4$, β taking on the values $b_1, \cdots, b_4$, and γ taking on the values $c_1, \cdots, c_4$. The marginal distribution in question is then defined as the sum of 4^7 terms:

$$\sum_{\alpha_1}\sum_{\alpha_2}\sum_{\alpha_3}\sum_{\beta_1}\sum_{\beta_2}\sum_{\gamma_1}\sum_{\gamma_2} p(a_1^2a_2\alpha_1\alpha_2\alpha_3b_1b_3^3\beta_1\beta_2c_1c_2c_4^2\gamma_1\gamma_2)$$
$$= p(a_1^2a_2b_1b_3^2c_1c_2c_4^2).$$

It is obvious that in the same way any marginal distribution for any m, r, s may be defined.

We return to the expression $S \cdot p'(a_{11}^{z_{11}}, \cdots, a_{mr}^{z_{mr}})$ considered in the first paragraph of this section, which we want to express in terms of products of the values of w' and of the s.d. According to the meaning of the three distributions involved, considering our hypotheses of random mating and of chromosome segregation, each w' will be multiplied by a sum of certain λ-values; or, the other way around, each λ-probability will be multiplied by a sum of certain w'-values. We want to find the coefficient of $\lambda(\varepsilon_1, \cdots, \varepsilon_m)$ in this expansion (see section 4 of I for the definition of $\lambda(\varepsilon_1, \cdots, \varepsilon_m)$) i.e., we want to collect all those w'-values which multipy a fixed $\lambda(\varepsilon_1, \cdots, \varepsilon_m)$ so as to have a formula

$$Sp'(a_{11}^{z_{11}} \cdots a_{1r}^{z_{1r}}, \cdots a_{m1}^{z_{m1}} \cdots a_{mr}^{z_{mr}})$$
$$= \sum_{\varepsilon_1 \ldots \varepsilon_m} \lambda(\varepsilon_1, \cdots \varepsilon_m) \cdot [\quad]. \tag{1}$$

Here, according to the meaning of the s.d., a $\lambda(\varepsilon_1, \cdots \varepsilon_m)$ is multiplied by a sum of w'-values each of which contains before the semicolon ε_i "correct values" corresponding to the ith locus and after the semicolon $(s-\varepsilon_i)$ "correct values" corresponding to this locus. By ε_i "correct values" we mean ε_i of *these* values which stand with respect to the ith locus on the left-hand side of the preceding equation which means ε_i values of the set $a_{i1}^{z_{i1}} \cdots a_{ir}^{z_{ir}}$ $(i = 1, \cdots m)$. Hence such a w' will be of the form

$$w'(a_{11}^{x_{11}} \cdots\cdot a_{1r}^{x_{1r}} \cdots\cdot, \cdots\cdot a_{m1}^{x_{m1}} \cdots\cdot a_{mr}^{x_{mr}} \cdots\cdot; \\ a_{11}^{y_{11}} \cdots\cdot a_{1r}^{y_{1r}} \cdots\cdot, a_{m1}^{y_{m1}} \cdots\cdot a_{mr}^{y_{mr}} \cdots\cdot),$$

where

$$\begin{gathered} x_{i1} + \cdots\cdot x_{ir} = \varepsilon_i, \qquad y_{i1} + \cdots\cdot y_{ir} = s - \varepsilon_i; \\ x_{ik} + y_{ik} = z_{ik}, \qquad z_{i1} + \cdots\cdot z_{ir} = s, \\ (i = 1, \cdots\cdot m), (k = 1, \cdots\cdot r). \end{gathered} \tag{2}$$

In such a w' there are still $(s - \varepsilon_1) + (s - \varepsilon_2) + \cdots\cdot + (s - \varepsilon_m)$ "free" or "empty" places before, and $\varepsilon_1 + \varepsilon_2 + \cdots\cdot + \varepsilon_m$ "free" or "empty" places after, the semicolon. Consider for the moment, to fix ideas, a fixed choice for the values on those $[(s - \varepsilon_1) + \varepsilon_1] + \cdots\cdot + [(s - \varepsilon_m) + \varepsilon_m] = ms$ empty places. Let, e.g., the first s of them all be equal to a_{11}, the second group all be equal to a_{21}, etc., and let us count the number of w' values which multiply $\lambda(\varepsilon_1, \cdots\cdot \varepsilon_m)$ for this particular choice and for a particular set of x_{ik}, y_{ik}. The ε_i "correct" values before the semicolon can be arranged in $\binom{s}{\varepsilon_i}$ ways on their s places and the $(s - \varepsilon_i)$ values after the semicolon in $\binom{s}{s - \varepsilon_i}$ ways on the s places available for them. Moreover, these ε_i values can still be arranged in $\dfrac{\varepsilon_i!}{x_{i1}! \cdots\cdot x_{ir}!}$ ways within the chosen ε_i places and, similarly, there are $\dfrac{(s - \varepsilon_i)!}{y_{i1}! \cdots\cdot y_{ir}!}$ possibilities for the $(s - \varepsilon_i)$ "correct" values after the semicolon. Hence the coefficient of a w' is

$$\prod_i^{1\ldots m} \left[\binom{s}{\varepsilon_i}\right]^2 \frac{\varepsilon_i!}{x_{i1}! \cdots\cdot x_{ir}!} \cdot \frac{(s - \varepsilon_i)!}{y_{i1}! \cdots\cdot y_{ir}!}.$$

Next, according to equation (4) of I replace each w' by the corresponding product of two p-values and divide both sides of (1) by $(s!)^m \prod\limits_i^{1\ldots m} \dfrac{1}{z_{i1}! \cdots\cdot z_{ir}!}$. The coefficient of $\lambda(\varepsilon_1, \cdots\cdot \varepsilon_m)$ in the expansion of $p'(a_{11}^{z_{11}} \cdots\cdot a_{mr}^{z_{mr}})$ is then a sum of terms of the form

$$\begin{gathered} \frac{1}{(s!)^m} \cdot \prod_i^{1\ldots m} \left[\binom{s}{\varepsilon_i}\binom{s}{\varepsilon_i} \frac{\varepsilon_i!}{x_{i1}! \cdots\cdot x_{ir}!} \cdot \frac{(s - \varepsilon_i)!}{y_{i1}! \cdots\cdot y_{ir}!} \cdot z_{i1}! \cdots\cdot z_{ir}!\right] \\ \cdot p(a_{11}^{x_{11}} \cdots\cdot a_{1r}^{x_{1r}} \cdots\cdot, \cdots\cdot, a_{m1}^{x_{m1}} \cdots\cdot a_{mr}^{x_{mr}} \cdots\cdot) \\ \cdot p(a_{11}^{y_{11}} \cdots\cdot a_{1r}^{y_{1r}} \cdots\cdot, \cdots\cdot, a_{m1}^{y_{m1}} \cdots\cdot a_{mr}^{y_{mr}} \cdots\cdot). \end{gathered} \tag{3}$$

So far we have assumed one fixed choice of values on the $(m \cdot s)$

"empty" places. Now we abandon this assumption. In the first p–probability in (3) there are $(s—\varepsilon_1)$ empty places after $a_{1r}^{x_{1r}}, \cdots\cdots,$ and $(s—\varepsilon_m)$ empty places after $a_{mr}^{x_{mr}}$. Similarly, there are ε_i, $(i = 1, \cdots\cdots m)$ empty places in the second p-probability. On the $(s—\varepsilon_i)$ free places in the first p–function we have to write the r values $a_{i1}, \cdots\cdots a_{ir}$ in all possible $r^{s-\varepsilon_i}$ combinations, and, similarly, for the second p–function. Then all these products of p–values are to be added. This amounts to an $[(s-\varepsilon_1)+\varepsilon_1]+\cdots\cdots+[(s-\varepsilon_m)+\varepsilon_m] = ms$–tuple sum of such products. But this sum of products obviously resolves into the product of two sums, one an $[(s-\varepsilon_1)+\cdots\cdots+(s-\varepsilon_m)]$–tuple sum, the other an $[\varepsilon_1+\cdots\cdots+\varepsilon_m]$–tuple sum, these sums being nothing else than the marginal distributions

$$p(a_{11}^{x_{11}}\cdots\cdots a_{1r}^{x_{1r}}, \cdots\cdots a_{m1}^{x_{m1}}\cdots\cdots a_{mr}^{x_{mr}})$$

and

$$p(a_{11}^{y_{11}}\cdots\cdots a_{1r}^{y_{1r}}, \cdots\cdots a_{m1}^{y_{m1}}\cdots\cdots a_{mr}^{y_{mr}})$$

of order $[\varepsilon_1+\cdots\cdots+\varepsilon_m]$ and $[(s-\varepsilon_1)+\cdots\cdots+(s-\varepsilon_m)]$ respectively. Thus, for a fixed set of x_{ik}, y_{ik} the coefficient of $\lambda(\varepsilon_1, \cdots\cdots \varepsilon_m)$ becomes

$$\prod_{i}^{1\ldots m}\left[\binom{s}{\varepsilon_i}\binom{s}{\varepsilon_i}\frac{\varepsilon_i!}{x_{i1}!\cdots\cdots x_{ir}!}\cdot\frac{(s-\varepsilon_i)!}{y_{i1}!\cdots\cdots y_{ir}!}\cdot\frac{z_{i1}!\cdots\cdots z_{ir}!}{s!}\right. \tag{4}$$

$$\cdot\, p(a_{11}^{x_{11}}\cdots\cdots a_{mr}\,x_{mr})\cdot p(a_{11}^{y_{11}}\cdots\cdots a_{mr}\,y_{mr}),$$

with (2) holding for the $x_{ik} \cdot y_{ik}, z_{ik}$.

This last expression can be simplified. Observing that $y_{ik} = z_{ik} - x_{ik}$ we have

$$\frac{s!}{\varepsilon_i!(s—\varepsilon_i)!}\cdot\frac{\varepsilon_i!}{x_{i1}!\cdots\cdots x_{ir}!}\cdot\frac{(s-\varepsilon_i)!}{y_{i1}!\cdots\cdots y_{ir}!}\cdot\frac{z_{i1}!\cdots\cdots z_{ir}!}{s!}$$

$$=\frac{z_{i1}!}{x_{i1}!(z_{i1}-x_{i1})!}\cdot\cdots\cdots\cdot\frac{z_{ir}!}{x_{ir}!(z_{ir}-x_{ir})!}.$$

Hence, introducing as before (section 4 of I)

$$l(\varepsilon_1, \cdots\cdots \varepsilon_m) = \binom{s}{\varepsilon_1}\cdots\cdots\binom{s}{\varepsilon_m}\lambda(\varepsilon_1, \cdots\cdots \varepsilon_m),$$

and using as always the notation $\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!(\alpha-\beta)!}$, $(0 \leqq \beta \leqq \alpha)$,

we arrive at the very simple and *general result*

$$p'(a_{11}^{z_{11}} \cdots\cdot a_{1r}^{z_{1r}}, \cdots\cdot a_{m1}^{z_{m1}} \cdots\cdot a_{mr}^{z_{mr}})$$

$$= \sum_{\varepsilon_1}^{0\cdots\cdot s} \cdots\cdot \sum_{\varepsilon_m}^{0\cdots\cdot s} l(\varepsilon_1, \cdots\cdot \varepsilon_m)$$

$$\cdot \sum_{x_{11}} \cdots\cdot \sum_{x_{1r}} \cdots\cdot \sum_{x_{m1}} \cdots\cdot \sum_{x_{mr}} \binom{z_{11}}{x_{11}} \cdots\cdot \binom{z_{1r}}{x_{1r}} \cdots\cdot \binom{z_{m1}}{x_{m1}} \cdots\cdot \binom{z_{mr}}{x_{mr}} \tag{5}$$

$$\cdot\, p(a_{11}^{x_{11}} \cdots\cdot a_{mr}^{x_{mr}}) \cdot p(a_{11}^{z_{11}-x_{11}} \cdots\cdot a_{mr}^{z_{mr}-x_{mr}}),$$

where

$$z_{i1} + \cdots\cdot z_{ir} = s, \qquad x_{i1} + \cdots\cdot x_{ir} = \varepsilon_i,$$

$$0 \leqq x_{ik} \leqq z_{ik}, \qquad (i = 1, \cdots\cdot m), \; (k = 1, \cdots\cdot r). \tag{6}$$

This is the general recurrence formula for the distributions of gametes in two successive generations for a $2s$-ploid with m loci and r alleles under random mating and chromosome segregation. It contains as particular cases the respective recurrence formulas of Geiringer (1944) and (1948a).

We finish this section by considering a few *particular cases.* First, let us consider a *homozygous organism.* We then put

$$a_{11} = \cdots\cdot = a_{1r} = A_1, \cdots\cdot, a_{m1} = \cdots\cdot a_{mr} = A_m,$$

and obtain the simple result

$$p'(A_1^s \cdots\cdot A_m^s) = \sum_{\varepsilon_1}^{0\cdots\cdot s} \cdots\cdot \sum_{\varepsilon_m}^{0\cdots\cdot s} l(\varepsilon_1 \cdots\cdot \varepsilon_s) \cdot \binom{s}{\varepsilon_1} \cdots\cdot \binom{s}{\varepsilon_m}$$

$$\cdot\, p(A_1^{\varepsilon_1} \cdots\cdot A_m^{\varepsilon_m}) \cdot p(A_1^{s-\varepsilon_1} \cdots\cdot A_m^{s-\varepsilon_m}). \tag{7}$$

The interpretation of (7) is obvious: $l(\varepsilon_1 \cdots\cdot \varepsilon_m) \cdot \binom{s}{\varepsilon_1} \cdots\cdot \binom{s}{\varepsilon_m}$ is the probability that ε_i maternal and $(s - \varepsilon_i)$ paternal genes will be transmitted with respect to the ith locus $(i = 1, \cdots\cdot m)$, while $p(A_1^{\varepsilon_1} \cdots\cdot A_m^{\varepsilon_m})$ is the probability of a maternal gamete with ε_i A_i-values, and $p(A_1^{s-\varepsilon_1} \cdots\cdot A_m^{s-\varepsilon_m})$ is the probability of a paternal one with $(s - \varepsilon_i)$ A_i-values, in the nth generation. Under random mating, the sum over all $\varepsilon_1, \cdots\cdot \varepsilon_m$ of the products of these three probabilities is the probability of a gamete of type $(A_1^s \cdots\cdot A_m^s)$ in the $(n+1)$st generation. It is not difficult to give an analogous interpretation for the general result contained in (6) (see Geiringer, 1948a, page 259). The biological situation considered in the present paper (general m, s, and r) is in a certain respect the most general one in which the recurrence formulas have the clear and simple structure exhibited in (6) or (7). If chromatid segregation is studied we find recurrence laws of an essentially different, much more complicated type, as will be shown by the author in a forthcoming paper.

Next consider the important case of just *two alleles*, $r=2$ at each locus, A_i and a_i $(i=1, \cdots m)$. Formula (6) then yields the result

$$\begin{aligned}
&p'(A_1^{z_1}a_1^{s-z_1}\cdots A_m^{z_m}a_m^{s-z_m}) \\
&= \sum_{\varepsilon_1}^{0\ldots s}\cdots\sum_{\varepsilon_m}^{0\ldots s} l(\varepsilon_1\cdots\varepsilon_m) \\
&\cdot\sum_{x_1}\cdots\sum_{x_m}\binom{z_1}{x_1}\binom{s-z_1}{\varepsilon_1-x_1}\cdots\binom{z_m}{x_m}\binom{s-z_m}{\varepsilon_m-x_m} \qquad (8)\\
&\cdot p(A_1^{x_1}a_1^{\varepsilon_1-x_1}\cdots A_m^{x_m}a_m^{\varepsilon_m-x_m}) \\
&\cdot p(A_1^{z_1-x_1}a_1^{s-z_1-\varepsilon_1+x_1}\cdots A_m^{z_m-x_m}a_m^{s-z_m-\varepsilon_m+x_m}).
\end{aligned}$$

In (8) the summation with respect to the x's is reduced to an m–tuple one rather than an $2m$–tuple one, as would follow from (6).

Finally, and this will give a very useful result, consider *the case where* r $\geqq$ s *and*

$$z_{11}=\cdots=z_{1r}=z_{21}=\cdots=z_{2r}=\cdots=z_{m1}=\cdots z_{mr}=1.$$

Without loss of generality we obtain the formula

$$\begin{aligned}
&p'(a_{11}\cdots a_{1s}\cdots a_{m1}\cdots a_{ms})=\sum_{\varepsilon_1}^{0\ldots s}\cdots\sum_{\varepsilon_m}^{0\ldots s} l(\varepsilon_1,\cdots\varepsilon_m) \\
&\cdot[p(a_{11}\cdots a_{1\varepsilon_1}\cdots a_{m1}\cdots a_{m,\varepsilon_m}) \qquad (9)\\
&\cdot p(a_{1,\varepsilon_1+1}\cdots a_{1s}\cdots a_{m,\varepsilon_m+1}\cdots a_{ms})+\cdots].
\end{aligned}$$

There are in each bracket $\binom{s}{\varepsilon_1}\cdots\binom{s}{\varepsilon_m}$ products of which the first has been written down. Let us however mention that the bracket which multiplies $l(\mu, \mu, \cdots \mu)$ contains $\frac{1}{2}\cdot\left[\binom{s}{\mu}\right]^m$ different terms only, while that whole bracket is multiplied by "two" [see (13)].

Let us show an illustration of (9). Consider $m=2$, $s=3$, write a_1, a_2, a_3 for the first three and b_1, b_2, b_3 for the other alleles. We find

$$\begin{aligned}
p'(a_1a_2a_3b_1b_2b_3) = {} & [l(00) + l(33)]\, p(a_1a_2a_3b_1b_2b_3) \\
& + [l(10) + l(23)] \cdot [p(a_1)p(a_2a_3b_1b_2b_3) + \cdots + \cdots] \\
& + [l(01) + l(32)] \cdot [p(b_1)p(a_1a_2a_3b_2b_3) + \cdots + \cdots] \\
& + [l(20) + l(13)] \cdot [p(a_1a_2)p(a_3b_1b_2b_3) + \cdots + \cdots] \\
& + [l(11) + l(22)] \cdot [p(a_1b_1)p(a_2a_3b_2b_3) \\
& \qquad + \cdots \text{ (nine terms)}] \\
& + [l(02) + l(31)] \cdot [p(b_1b_2)p(a_1a_2a_3b_3) + \cdots + \cdots] \\
& + [l(30) + l(03)]\, p(a_1a_2a_3)p(b_1b_2b_3) \\
& + l(12) + l(21)] \cdot [p(a_1b_1b_2)p(a_2a_3b_3) \\
& \qquad + \cdots \text{ (nine terms)}].
\end{aligned} \tag{10}$$

Next consider the case $m=2$, $s=4$. We want to write down only the term which multiplies $2l(22)$. It is:

$$\begin{aligned}
2l(22)\, [& p(a_1a_2b_1b_2)p(a_3a_4b_3b_4) + p(a_1a_2b_1b_3) \\
& \cdot p(a_3a_4b_2b_4) + p(a_1a_2b_1b_4)p(a_3a_4b_2b_3) \\
& + \cdots \text{ (eighteen terms)}].
\end{aligned} \tag{11}$$

Since $l(\varepsilon_1, \cdots \varepsilon_m) = l(s - \varepsilon_1, \cdots s - \varepsilon_m)$ the formulas for $s=2\mu$ as well as for $s=2\mu + 1$ take on the following form

$$\begin{aligned}
& p'(a_{11} \cdots a_{1s} \cdots a_{m1} \cdots a_{ms}) = 2 \sum_{\varepsilon_1 \ldots \varepsilon_m}^{0 \ldots \mu}{}' \, l(\varepsilon_1 \cdots \varepsilon_m) \\
& \cdot [p(a_{11} \cdots a_{1\varepsilon_1} \cdots a_{m1} \cdots a_{m,\varepsilon_m}) \\
& \cdot p(a_{1,\varepsilon_1+1} \cdots a_{1s} \cdots a_{m,\varepsilon_m+1} \cdots a_{ms}) + \cdots].
\end{aligned} \tag{12}$$

Here $\sum'$ indicates that for $s=2\mu$ the expression in the brackets which multiplies $2l(\mu, \mu, \cdots \mu)$ contains $\frac{1}{2} \cdot \left[\binom{2\mu}{\mu}\right]^m$ terms only. As an example consider for $m=3$, $s=2$:

$$\begin{aligned}
p'(a_1a_2b_1b_2c_1c_2) = {} & 2l(000)p(a_1a_2b_1b_2c_1c_2) \\
& + 2l(100)[p(a_1)p(a_2b_1b_2c_1c_2) + \cdots] + \cdots + \cdots \\
& + 2l(110)[p(a_1b_1)p(a_2b_2c_1c_2) + \cdots + \cdots + \cdots] \\
& \qquad + \cdots + \cdots \\
& + 2l(111)[p(a_1b_1c_1)p(a_2b_2c_2) + \cdots + \cdots + \cdots].
\end{aligned} \tag{13}$$

From (9) or (12) we can now find any formula regarding partly equal alleles by equating some of the a_{1k} to each other, some of the a_{2k} to each other, etc. If for example in (11), $a_1=a_2=a_3=a_4=a$ the coefficient of $l(22)$ will be:

$$\begin{aligned}
12\, l(22) \cdot [& p(a^2b_1b_2)p(a^2b_3b_4) + p(a^2b_1b_3)p(a^2b_2b_4) \\
& + p(a^2b_1b_4)p(a^2b_2b_3)].
\end{aligned}$$

It is thus easy to write down any particular recurrence formula. By means of such a formula the distributions of gametes (and therefore because of equation (4) of I the distributions of zygotes) can be found for any generation if the initial distribution is known.

The recurrence formulas are, mathematically, *quadratic difference equations whose constant coefficients are the values of our s.d.* The decisive role which the s.d. plays in the organization of our *biologically meaningful* recurrence relations seems to hint that our concept of a s.d. is not inadequate biologically. On the other hand, once our recurrence formulas have been established, they are completely independent of any particular assumptions regarding the parameters in the s.d. Random chromosome segregation or many other models of segregation may be used (see also section 4 of I). No recurrence formulas have been given by other authors concerned with our problem.

2. *Recurrence Formulas for the Marginal Distributions.*

If we wish to *integrate* the recurrence formulas of the preceding section or if we want to study the *limit behavior of our distributions,* we need recurrence formulas for the marginal distributions. They can be derived from the recurrence relations of section 1 of II (see also section 5 of I). Consider first the case $m=s=r=2$ where:

$$\begin{aligned} p'(a_1a_2b_1b_2) = {} & 2l(00)p(a_1a_2b_1b_2) + 2l(01) \\ & \cdot [p(b_1)p(a_1a_2b_2) + p(b_2)p(a_1a_2b_1)] + 2l(10) \\ & \cdot [p(a_1)p(a_2b_1b_2) + p(a_2)p(a_1b_1b_2)] \\ & + 2l(02)p(a_1a_2)p(b_1b_2) + 2l(11)[p(a_1b_1)p(a_2b_2) \\ & + p(a_1b_2)p(a_2b_1)], \end{aligned}$$

and

$$\begin{aligned} p'(a_1a_2b_1^2) = {} & 2l(00)p(a_1a_2b_1^2) + 4l(01)p(b_1)p(a_1a_2b_1) \\ & + 2l(10)[p(a_1)p(a_2b_1^2) + p(a_2)p(a_1b_1^2)] \\ & + 2l(02)p(a_1a_2)p(b_1^2) + 4l(11)p(a_1b_1)p(a_2b_1). \end{aligned}$$

Upon addition we obtain

$$\begin{aligned} p'(a_1a_2b_1) & \equiv p'(a_1a_2b_1b_2) + p'(a_1a_2b_1b_1) \\ & = [2l(00) + 2l(01)]p(a_1a_2b_1) \\ & + [2l(01) + 2l(02)]p(a_1a_2)p(b_1) \\ & + [2l(10) + 2l(11)][p(a_1)p(a_2b_1) + p(a_2)p(a_1b_1)]. \end{aligned}$$

Here the coefficients of $p(a_1a_2b_1)$ or of $p(a_1a_2)$ $p(b_1)$ to the right are, of course, *the marginal distributions of our s.d.* as introduced in section 4 of I. Using the notation of section 4 of I we have to write

$$p'(a_1a_2b_1) = 2l^{(0)(1)}(00)p(a_1a_2b_1) + 2l^{(0)(1)}(01)p(a_1a_2)p(b_1) \\ + 2l^{(0)(1)}(10)[p(a_1)p(a_2b_1) + p(a_2)p(a_1b_1)], \qquad (14)$$

with the symmetry properties as in equation (36) of I

$$l^{(0)(1)}(00) = l^{(0)(1)}(21), \quad l^{(0)(1)}(10) = l^{(0)(1)}(11), \\ l^{(0)(1)}(01) = l^{(0)(1)}(20).$$

We may derive a recurrence formula for $p'(a_1b_1b_2)$ in the same way. Next, by a twofold summation, still for $s=m=2$ we obtain

$$p'(a_1a_2) = 2l^{(0)(2)}(00)p(a_1a_2) + 2l^{(0)(2)}(10)p(a_1)p(a_2), \qquad (15)$$

where

$$l^{(0)(2)}(00) = l(00) + 2l(01) + l(02); \\ l^{(0)(2)}(10) = l(10) + 2l(11) + l(12),$$

and

$$p'(a_1b_1) = 2l^{(1)(1)}(00)p(a_1b_1) + 2l^{(1)(1)}(10)p(a_1)p(b_1), \qquad (16)$$

with

$$l^{(1)(1)}(00) = l(00) + l(10) + l(01) + l(11); \\ l^{(1)(1)}(10) = l(10) + l(20) + l(11) + l(21);$$

and, of course, in a similar manner we obtain the recurrence relations for $p'(b_1b_2)$, for $p'(a_1)$, and for $p'(b_1)$.

It does not seem necessary to show the marginal character of the distribution to the left in an explicit way in the notation, since this becomes quite clear by the superscripts in the values of the s.d. to the right. If, for example, the distribution to the left is $p'(a_1b_1)$ while on the right there is a term containing $l^{(1)(1)}(00)$, it is clear that we deal with an $s=m=2$-organism. Or, if in a formula where $p'(a_1a_2b_1b_2c_1)$ stands to the left, on the right there is a term containing, for example, $l^{(3)(3)(4)(5)}(0000)$ we see at once that the organism considered is of type $m=4$, $s=5$ since $2+3=2+3=1+4=0+5=5$.

Using the concepts introduced in section 4 of I and the results of section 1 of II we obtain easily the recurrence formula for any marginal distribution. Consider the case just mentioned and write now for the sake of brevity $l^{(3)(3)(4)(5)} = \bar{l}$. We have

$$
\begin{aligned}
p'(a_1a_2b_1b_2c_1) = {} & 2\bar{l}(0000)\,p(a_1a_2b_1b_2c_1) \\
& + 2\bar{l}(1000)\cdot[p(a_1)\,p(a_2b_1b_2c_1) + p(a_2)\,p(a_1b_1b_2c_1)] \\
& + 2\bar{l}(0100)\cdot[p(b_1)\,p(a_1a_2b_2c_1) + p(b_2)\,p(a_1a_2b_1c_1)] \\
& + 2\bar{l}(0010)\,p(c_1)\,p(a_1a_2b_1b_2) \\
& + 2\bar{l}(1100)\,[p(a_1b_1)\,p(a_2b_2c_1) + \cdots + \cdots + \cdots] + \cdots + \cdots \\
& + 2\bar{l}(2000)\,p(a_1a_2)\,p(b_1b_2c_1) + 2\bar{l}(0200)\,p(b_1b_2)\,p(a_1a_2c_1).
\end{aligned}
\tag{17}
$$

If in (17) we then put $a_1 = a_2$ we find a formula for $p'(a_1^2b_1b_2c_1)$.

A general formula for the marginal distributions, analogous to (12), is as follows. Put $s - \nu_i = 2\mu_i$, or $s - \nu_i = 2\mu_i + 1$ respectively ($i = 1, \cdots m$). We obtain

$$
\begin{aligned}
& p'(a_{11} \cdots a_{1,s-\nu_1} \cdots a_{m1} \cdots a_{m,s-\nu_m}) \\
& = 2 \sum_{\varepsilon_1}^{0 \ldots \mu_1} \cdots \sum_{\varepsilon_m}^{0 \ldots \mu_m} l^{(\nu_1) \ldots (\nu_m)}(\varepsilon_1, \cdots \varepsilon_m) \\
& \cdot [p(a_{11} \cdots a_{1,\varepsilon_1} \cdots a_{m1} \cdots a_{m,\varepsilon_m}) \\
& \cdot p(a_{1,\varepsilon_1+1} \cdots a_{1,s-\nu_1} \cdots a_{m,\varepsilon_m+1} \cdots a_{m,s-\nu_m}) + \cdots].
\end{aligned}
\tag{18}
$$

The recurrence formulas (12)—(18) are quite general and seem to reveal a regularity of the heredity mechanism. They are so simple that we may write them down for any m, r, and s without computation. The values of the s.d., or of the marginal distributions of the s.d., appear each time as the constant coefficient in these quadratic recurrence relations (finite difference equations). The essential concepts are: (1) the s.d., and (2) its marginal distributions; (3), the d.ga., and (4) its marginal distributions. Without the use of these concepts the computations become hopelessly involved, as exemplified in the explicit computation of section 5 of I. Recurrence formulas for the simplest cases are due to Jennings for $s=1$, $m=r=2$, and to Haldane for $m=1$, $s=r=2$. The essential restrictions for our recurrence laws are: (1) random mating, (2) chromosome segregation rather than chromatid segregation, and (3) the s.d. independent of the genotype of the organism and of "n." (It may be noted that the essential structure of the recurrence laws does not change if different s.d.'s are assumed for males and for females.)

It is not difficult to work out an *integration* of our recurrence relations and to arrive at explicit formulas in a few simple cases. By that we mean, for example, that $p^{(n)}(a_1a_2b_1b_2)$ is expressed in terms of $p^{(0)}(a_1a_2b_1b_2)$, of the marginal distributions of $p^{(0)}$, of the parameters introduced in the definition of the s.d., and of course, of "n." The explicit formulas for $m=2$, $s=1$ and $s=2$, $m=1$ are due to Jennings and Haldane, and for $s=1$, $m=3$ to the author, who worked out a general procedure for $s=1$, any m, and for $m=1$, any s, (see Geiringer, 1944, section 8; 1948a, section 6).

3. *The Limit Theorem for* m=s=2.

We begin with the consideration of the case $m=s=2$ where all possibilities can be discussed explicitly. This discussion will be needed in the treatment of the general case.

1) With the notations of section 2 of I we first assume

$$l_1(0) \equiv l(00) + 2l(01) + l(02) < \tfrac{1}{2}; \tag{19}$$

$$l_2(0) \equiv l(00) + 2l(10) + l(20) < \tfrac{1}{2}. \tag{20}$$

It follows that

$$l_{12}(00) \equiv l(00) + l(10) + l(01) + l(11) < \tfrac{1}{2}. \tag{21}$$

In fact, if $2l_{12}(00) = 2l(00) + 2l(10) + 2l(01) + 2l(11) = 1$ would hold together with

$$2l(00) + 4l(01) + 4l(10) + 4l(11) + 2l(20) = 1, \tag{22}$$

it would follow that $l(10) = l(01) = l(20) = l(11) = 0$ and $l(00) = 1/2$, and then (19) and (20) could not be true. We formulate this intermediate result explicitly:

If $\mathrm{l}_{12}(00) = 1/2$, $\mathrm{c}_{12} = \mathrm{l}_{12}(10) + \mathrm{l}_{12}(01) = 0$, *then* $\mathrm{l}_1(0) = \mathrm{l}_2(0) = 1/2$. *Hence if at least one of the two inequalities* (19)—(20) *holds,* $\mathrm{c}_{12} > 0$.

Now consider the recurrence formulas. Writing $p^{(n)} = p$, $p^{(n+1)} = p'$, $a_1 = A$, $a_2 = a$, $b_1 = B$, $b_2 = b$:

$$\begin{aligned} p'(AaBb) = {} & 2l(00)p(AaBb) + 2l(10)\cdot[p(A)p(aBb) \\ & + p(a)p(ABb)] + 2l(01)[p(B)p(AaB) + p(b)p(AaB)] \\ & + 2l(20)p(Aa)p(Bb) \\ & + 2l(11)[p(AB)p(ab) + p(Ab)p(aB)], \end{aligned} \tag{23}$$

and

$$\begin{aligned} p'(AaB) = {} & 2l^{(0)(1)}(00)p(AaB) + 2l^{(0)(1)}(01)p(Aa)p(B) \\ & + 2l^{(0)(1)}(10)[p(A)p(aB) + p(a)p(AB)]; \end{aligned} \tag{24}$$

$$p'(AB) = 2l^{(1)(1)}(00)p(AB) + 2l^{(1)(1)}(10)p(A)p(B); \tag{25}$$

$$p'(Aa) = 2l^{(0)(2)}(00)p(Aa) + 2l^{(0)(2)}(10)p(A)p(a). \tag{26}$$

To evaluate them we shall need, as in previous papers, the

Lemma A). *If* $\mathrm{x}_{n+1} = \alpha \mathrm{x}_n + \mathrm{y}_n$ $(n=0, 1, \cdots)$ *where* $|\alpha| < 1$, *and if* $\lim_{n\to\infty} \mathrm{y}_n = \mathrm{y}$, *then* $\lim_{n\to\infty} \mathrm{x}_n = \mathrm{y}/(1-\alpha)$.

We now add the recurrence formulas for $p'(A^2)$ and $p'(Aa)$ using (26). The result is that

$$p^{(n+1)}(A) = p^{(n)}(A),$$

and there is an analogous result for $p^{(n)}(B)$. Thus

$$p^{(n)}(A) = p^{(0)}(A),\ p^{(n)}(a) = p^{(0)}(a),$$
$$p^{(n)}(B) = p^{(0)}(B), p^{(n)}(b) = p^{(0)}(b),\ (n=0, 1, \cdots).$$

Next we apply Lemma A) to (26). Here $\alpha = 2l^{(0)(2)}(00) = 2l_1(0) < 1$ because of (19). Hence, considering that $2l^{(0)(2)}(00) + 2l^{(0)(2)}(10) = 1$, it follows from Lemma A) that

$$\lim_{n\to\infty} p^{(n)}(Aa) = p^{(0)}(A)p^{(0)}(a),$$

and in the same way, since $2l_2(0) < 1$, it follows that

$$\lim_{n\to\infty} p^{(n)}(Bb) = p^{(0)}(B)p^{(0)}(b).$$

Also, since $2l^{(1)(1)}(00) = 2l_{12}(00) < 1$ it follows that

$$\lim_{n\to\infty} p^{(n)}(AB) = p(A)p(B).$$

Next consider (24) and use now $p^{(n)}, p^{(n+1)}$, and $p^{(0)} \equiv p$:

$$\begin{aligned} &p^{(n+1)}(AaB) - 2l^{(0)(1)}(00)p^{(n)}(AaB) \\ &= 2l^{(0)(1)}(01)p^{n}(Aa)p(B) + 2l^{(0)(1)}(10) \cdot [p(A)p^{(n)}(aB) \\ &\qquad + p(a)p^{(n)}(AB)]. \end{aligned}$$

As $n \to \infty$ the right side converges toward

$$\begin{aligned} &p(A)p(a)p(B)[2l^{(0)(1)}(10) + 2l^{(0)(1)}(01)] \\ &= [1 - 2l^{(0)(1)}(00)]p(A)p(a)p(B). \end{aligned}$$

Now

$$l^{(0)(1)}(00) = l(00) + l(01) < l_1(0) < \tfrac{1}{2}.$$

Hence the conditions of Lemma A) are again satisfied and

$$\lim_{n\to\infty} p^{(n)}(AaB) = p(A)p(a)p(B)$$

and an analogous conclusion holds for $p^{(n)}(ABb)$, etc. Finally since $2l(00) < 1$, we obtain the following result which covers the "general case" in which $l_1(0) < 1/2$, $l_2(0) < 1/2$.

If (19) holds, then

$$\lim_{n\to\infty} p^{(n)}(AaBb) = p(A)p(a)p(B)p(b). \tag{27}$$

2) Next assume $l_1(0) = 1/2$, $l_2(0) < 1/2$. From the definitions of $l_1(0)$, $l_2(0)$, and $l_{12}(00)$ together with equation (9) of I it follows that in this case $l(10) = l(11) = 0$, $l(00) + l(20) < 1/2$, and $l_{12}(00) < 1/2$. The recurrence relation (23) then becomes

$$\begin{aligned} p^{(n+1)}(AaBb) = {} & 2l(00)p^{(n)}(AaBb) \\ & + 2l(01)[p^{(n)}(Aab)p(B) + p^{(n)}(AaB)p(b)] \\ & + 2l(02)p^{(n)}(Aa)p^{(n)}(Bb). \end{aligned} \tag{28}$$

Since $l_1(0) = 1/2$ it follows that $p^{(n)}(Aa) = p^{(0)}(Aa) \equiv p(Aa)$ $(n=0, 1, \cdots)$. In (28) $l(01) \neq 0$ and $2l(00) < 1$. In the relation for $p^{(n+1)}(AaB)$ the coefficient of $p^{(n)}(AaB)$ is $2l(00) + 2l(01)$ and this is, because of (9), less than one. Hence we have with $p^{(0)} = p$:

$$\lim_{n\to\infty} p^{(n)}(AaBb) = p(Aa)p(B)p(b).$$

We reason in the same way if $l_1(0) < 1/2$, $l_2(0) = 1/2$. Finally if $l_1(0) = l_2(0) = 1/2$, and consequently $l_1(1) = l_2(1) = 0$, it follows that $l(10) = l(01) = l(11) = 0$ and $2l(00) + 2l(02) = 1$. If these last two values are both different from zero the recurrence formula for $p^{(n+1)}(AaBb)$ together with $p^{(n)}(Aa) = p(Aa)$, $p^{(n)}(Bb) = p(Bb)$, $(n=0, 1, \cdots)$ shows that $p^{(n)}(AaBb) \to p(Aa)p(Bb)$. If $l(20) = 1/2$, $l(00) = 0$ the same limit is reached in the first filial generation, i.e. $p^{(n)}(AaBb) = p(Aa)p(Bb)$ $(n=0, 1, \cdots)$. If, finally, $2l(00) = 1$ we have complete linkage, and $p^{(n)}(AaBb) = p(AaBb)$ for $n=0, 1, \cdots$. Thence the theorem:

If, for m=s=2 *we have, in the "general case:"*

$$\begin{aligned} l_1(0) &= l(00) + 2l(01) + l(02) < \tfrac{1}{2}; \\ l_2(0) &= l(00) + 2l(10) + l(20) < \tfrac{1}{2}; \end{aligned}$$

then, writing $\mathrm{p}^{(0)} = \mathrm{p}$ *we have for each allele*

$$p^{(n)}(A) = p(A), \cdots, p^{(n)}(B) = p(B), \cdots \quad (n=0, 1, \cdots); \tag{29}$$

$$\lim_{n\to\infty} p^{(n)}(AaBb) = p(A)p(a)p(B)p(b). \tag{30}$$

If $\mathrm{l}_1(0) = 1/2$, $\mathrm{l}_2(0) < 1/2$, *then*

$$p^{(n)}(Aa) = p(Aa),\ p^{(n)}(B) = p(B),\ p^{(n)}(b) = p(b), \qquad (n=0,1,\cdots); \tag{31}$$

$$\lim_{n\to\infty} p^{(n)}(AaBb) = p(Aa)p(B)p(b), \tag{32}$$

and similar results hold if $\mathrm{l}_1(0) < 1/2$, $\mathrm{l}_2(0) = 1/2$.

If $\mathrm{l}_1(0) = \mathrm{l}_2(0) = 1/2$, *but* $\mathrm{l}(00) < 1/2$, *then*

$$p^{(n)}(Aa) = p(Aa), \quad p^{(n)}(Bb) = p(Bb), \quad (n=0,1,\cdots); \tag{33}$$

$$\lim_{n\to\infty} p^{(n)}(AaBb) = p(Aa)p(Bb). \tag{34}$$

Finally, if $\mathrm{l}(00) = 1/2$, *all other* l–*values are equal to zero and, in this case of complete linkage*:

$$p^{(n)}(AaBb) = p(AaBb), \quad (n=0,1,\cdots). \tag{35}$$

The corresponding limit results regarding the marginal distributions $p^{(n)}(AB)$, and $p^{(n)}(Aa)$ of our distribution $p^{(n)}(AaBb)$ are the well known limit results of Haldane and Jennings. For Haldane's s.d. (see section 3 of I) $l_1(0) = l_2(0) = 1/6$, hence the conditions (19) and (20) are satisfied and *the limit result* (30) *thus holds for Haldane's linkage theory.*

4. *The Main Limit Theorem.* (*General Case.*)

We shall consider in this section the "general case" which corresponds to (30) and to the situation of the main theorem of Geiringer (1944). A complete theorem covering the various cases of "degeneration" of the s.d. (analogous to that in Geiringer, 1945) will then be established in the last section.

For the sake of brevity we shall speak here of a "mixed" gamete (with respect to a certain locus) if with respect to this locus the gamete contains paternal as well as maternal heritage. The decisive condition in the limit theorem of Geiringer (1948a) ($m=1$, general "s") was that $l(0) < 1/2$, i.e., that the probability $1\text{–}2\,l(0)$ of mixed gametes be different from zero, while the decisive condition in the main theorem of Geiringer (1944) ($s=1$, general "m") was that all recombination probabilities (r.p.) $c_{i\gamma} = 1 - 2l_{i\gamma}(00)$ be different from zero. Let us write these conditions in the notations of the present paper where they apply to certain marginal distributions of our s.d. Remembering (37) and (39) of I we first consider

$$l^{(s-1)\ldots.(s-1)}(\varepsilon_1,\cdots\varepsilon_m) = l_{12\ldots.m}(\varepsilon_1,\cdots\varepsilon_m); \qquad (\varepsilon_i=0,1),\ (i=1,\cdots,m). \tag{36}$$

This is an ordinary linkage distribution as in Geiringer (1944). From $l_{12\ldots.m}(\varepsilon_1, \cdots\cdot \varepsilon_m)$ we derive in the usual way by $(m-2)$-tuple summation the marginal distribution $l_{ik}(\varepsilon_i \varepsilon_k)$. In our present notation this reads, for $i=1$, $k=2$,

$$l_{12}(\varepsilon_1 \varepsilon_2) = l^{(s-1)(s-1)(s)\ldots.(s)}(\varepsilon_1\, \varepsilon_2\, 0 \cdots\cdot 0) \equiv l_{12}{}^{(s-1)(s-1)}(\varepsilon_1\, \varepsilon_2)\,; \qquad (\varepsilon_1 = 0, 1, \varepsilon_2 = 0, 1), \tag{37}$$

and there are $m(m-1)/2$ such distributions. We also have

$$l^{(0)(s)\ldots.(s)}(\varepsilon_1, 0 \cdots\cdot 0) = l_1(\varepsilon_1), \quad (\varepsilon_1 = 0, 1, \cdots\cdot s),$$

and there are m such distributions.

It will be necessary that

$$l_{i\gamma}(00) < \tfrac{1}{2}; \quad (i, \gamma = 1, \cdots\cdot m), \tag{38}$$

and

$$l_i(0) < \tfrac{1}{2}; \quad (i = 1, \cdots\cdot m) \tag{39}$$

be satisfied. We shall see that these obviously necessary conditions are sufficient as well. Moreover we shall see (Lemma D) that (38) is a consequence of (39). It will be convenient to introduce a short name for $1-2l_i(0)$. In analogy to $1-2l_{ik}(00) = c_{ik}$ where c_{ik} is the r.p., put $1-2l_i(0) = c_i$ and call c_i "exchange probability." Then (39) states that all exchange probabilities are greater than zero.

Besides Lemma A) of the preceding section we shall need three more Lemmas:

Lemma B). *With the definition* (39) *of I we have for* $\nu = 0, 1, \cdots\cdot (s-2)$:

$$l^{(\nu)(s)\ldots.(s)}(\varepsilon_1) \equiv l_1{}^{(\nu)}(\varepsilon_1) = \tfrac{1}{2}, \quad (\varepsilon_1 = 0, 1, \cdots\cdot (s-\nu))$$

if, and only if,

$$l^{(0)(s)\ldots.(s)} \equiv l_1(\varepsilon_1) = \tfrac{1}{2}, \quad (\varepsilon_1 = 0, 1 \cdots\cdot s).$$

In fact for $m=1$ this is the Lemma B) of Geiringer (1948). For general m it is no new theorem either, since $l^{(0)(s)\ldots.(s)}(\varepsilon_1)$ as well as $l^{(\nu)(s)\ldots.(s)}(\varepsilon_1)$ are nothing but s.d.'s for $m=1$ and general "s." The equation which corresponds to equation (38) in Geiringer (1948) now reads

$$l_i{}^{(s-1)}(\varepsilon_i) \equiv l_i{}^{(s-1)(s)\ldots.(s)}(\varepsilon_i) = \tfrac{1}{2}, \quad (\varepsilon_i = 0, 1). \tag{40}$$

Next we need the:

Lemma C). *The inequality*

$$l^{(i)(\gamma)\cdots\cdot(k)}(0\,0\cdots\cdot 0) \leqq l^{(i')(\gamma')\cdots\cdot(k')}(0\,0\cdots\cdot 0) \tag{41}$$

holds if

$$i' \geqq i\,, \quad j' \geqq j\,, \cdots\cdot k' \geqq k\,. \tag{42}$$

This is almost obvious, since all terms contained in $l^{(i)(j)\cdots\cdot(k)}(0\,0\cdots\cdot 0)$ are contained in $l^{(i')(j')\cdots\cdot(k')}(0\,0\cdots\cdot 0)$ as well, and if at least one of the inequalities (42) holds the second expression will contain terms not contained in the first. (Consider as an example $l^{(1)(2)(1)}(000)$ compared with $l^{(1)(2)(2)}(000)$ which are given, symbolically by the expressions $(l_0 + l_1)(l_0 + 2l_1 + l_2)(l_0 + l_1)$ and $(l_0 + l_1)(l_0 + 2l_1 + l_2)(l_0 + 2l_1 + l_2)$ respectively.)

Let us return to the general recurrence formula (9) and the corresponding formulas for the marginal distributions which occur on the right-hand side of (9). Consider first all marginal distributions of order two, such as (avoiding double subscripts) $p^{(n)}(A^2)$, $p^{(n)}(Aa), \cdots\cdot p^{(n)}(AB)$, etc. We have with

$$l^{(s-2)(s)\cdots\cdot(s)}(\varepsilon) = l_1^{(s-2)}(\varepsilon):$$

$$p^{(n+1)}(A^2) = 2l_1^{(s-2)}(0)p^{(n)}(A^2) + 2l_1^{(s-2)}(1)[p(A)]^2\,.$$

Here, because of (39) and Lemma B) the coefficient of $p^{(n)}(A^2)$ is less than one, hence, because of Lemma A), $p^{(n)}(A^2) \rightarrow p(A)^2$. We also have

$$p^{(n+1)}(AB) + 2l_{12}^{(s-1)(s-1)}(00)p^{(n)}(AB)$$
$$+ 2l_{12}^{(s-1)(s-1)}(10)p(A)p(B).$$

Here $2l_{12}^{(s-1)(s-1)}(00) < 1$ because of (38). Using Lemma A) we see that $p^{(n)}(AB) \rightarrow p(A)p(B)$.

Next we consider the recurrence formulas for the marginal distributions of order three, such as $p^{(n)}(A^3)$, $\cdots\cdot$, $p^{(n)}(A^2B)$, $\cdots\cdot$, $p^{(n)}(ABC)$, $\cdots\cdot$. For $p^{(n)}(A^3)$ the convergence toward $[p(A)]^3$ follows because of (39). We also have

$$p^{(n+1)}(A^2B) = l_{12}^{(s-2)(s-1)}(00)p^{(n)}(A^2B) + \cdots\cdot.$$

Now $l_{12}^{(s-2)(s-1)}(00) < l_{12}^{(s-1)(s-1)}(00)$ because of Lemma C), and this last value is less than 1/2 according to (38). Hence because of Lemma A), and of the convergence of the distributions of order two we see that $p^{(n)}(A^2B) \rightarrow p(A)^2p(B)$. Also in

$$p^{(n+1)}(ABC) = 2l_{123}^{(s-1)(s-1)(s-1)}(000)p^{(n)}(ABC) + \cdots\cdot$$

we have $l_{123}^{(s-1)(s-1)(s-1)}(000) < l_{123}^{(s-1)(s-1)(s)}(000) = l_{12}^{(s-1)(s-1)}(00)$ and since this last one is less than 1/2, the coefficient of $p^{(n)}(ABC)$ is less than one and by means of Lemma A), it follows that $p^{(n)}(ABC) \rightarrow p(A)p(B)p(C)$.

We continue in the same way. The coefficient of $p^{(n)}(A^{\alpha})$ in the recurrence formula for $p^{(n+1)}(A^{\alpha})$ is $2l_1^{(s-\alpha)}(0)$ and this is less than one because of Lemma B) and (39), and the coefficient of a term like $p^{(n)}(A^{\alpha}B^{\beta}C^{\gamma})$ in the respective formula is

$$2l_{123}^{(s-\alpha)(s-\beta)(s-\gamma)(s)\cdots\cdots(s)}(000) \leqq 2l_{12}^{(s-1)(s-1)}(00) < 1,$$

because of Lemma C) and (38). Hence we may each time apply Lemma A) and the convergence toward the product of all marginal distributions of first order follows.

Before formulating a final result we prove:

Lemma D). *If a recombination probability* $\mathrm{c}_{ij} = 1\text{-}2\mathrm{l}_{ij}(00)$ *vanishes, then the two corresponding exchange probabilities* $\mathrm{c}_i = 1\text{-}2\mathrm{l}_i(0)$, *and* $\mathrm{c}_j = 1\text{-}2\mathrm{l}_j(0)$ *are zero as well.*

First let us prove this for $m=2$ and general "s." We have, symbolically,

$$2l_{12}(00) \sim 2[l_0 + (s-1)l_1 + \cdots\cdots + l_{s-1}]^2.$$

Now if

$$2l_{12}(00) = 1,$$

and since the sum of all l-values equals one,

$$1 \sim [l_0 + sl_1 + \cdots\cdots + l_s]^2,$$

we see, considering the symmetry relations for the s.d., that the coefficient of each single term to the right in the second of these equivalence relations is greater than of the corresponding term in the first one with the only exception of $l(00)$. Hence, if we subtract the smaller sum from the larger one it follows that each l-value equals zero except $l(00)$. But if $2l(00) = 1$, $c_{12} = 0$, then $l_1(0) = l_2(0) = 1/2$, and this proves our statement.

Now consider general m and s. Corresponding to the ith and jth factor we form the marginal distribution of order two for general s. For $i=1$, $j=2$ this gives

$$l^{(0)(0)(s)\cdots\cdots(s)}(\varepsilon_1\,\varepsilon_2\,0\cdots\cdots 0) = l_{12}^{(0)(0)}(\varepsilon_1\,\varepsilon_2), \quad (\varepsilon_i = 0, 1, \cdots\cdots s).$$

This distribution has all the properties of a s.d. for $m=2$, and general s. Hence for this distribution Lemma D) holds, which means if $l_{12}^{(s-1)(s-1)}(00) = l_{12}(00) = \frac{1}{2}$, then $l_1(0) = l_2(0) = 1/2$. Here

$l_{12}^{(s-1)(s-1)}(\varepsilon_1\varepsilon_2) = l_{12}(\varepsilon_1\varepsilon_2)$ has been derived from $l_{12}^{(0)(0)}(\varepsilon_1\varepsilon_2)$ by an $2 \cdot (s-1)$-tuple summation, and $l_1(\varepsilon_1)$ from the same distribution by an s-tuple summation. But these same distributions are also the marginal distributions of our original s.d. with arbitrary m. In fact

$$\begin{aligned} l_{12}(\varepsilon_1\ \varepsilon_2) &= l^{(s-1)(s-1)(s)\cdots\cdots(s)}(\varepsilon_1\ \varepsilon_2\ 0 \cdots\cdots 0), \quad (\varepsilon_1 = 0, 1, \varepsilon_2 = 0, 1); \\ l_1(\varepsilon_1) &= l^{(0)(s)\cdots\cdots(s)}(\varepsilon_1, 0 \cdots\cdots 0), \quad (\varepsilon_1 = 0, 1, \cdots\cdots s); \\ l_2(\varepsilon_2) &= l^{(s)(0)\cdots\cdots(s)}(0, \varepsilon_2, \cdots\cdots 0), \quad (\varepsilon_2 = 0, 1, \cdots\cdots s). \end{aligned}$$

Hence the Lemma is proved for the subscripts 1 and 2 and therefore for any pair of subscripts.

Let us note that conditions (39) are not only sufficient but necessary as well, considering the limit result we have in mind. This will be clarified in the following section. However, using this fact in the formulation of our theorem we obtain the following general result:

Theroem I. If, and only if, conditions (39) *hold, or in words, if and only if, the* m *exchange probabilities* $c_i = 1\text{-}2l_i(0)$ *are positive the distribution of gametes converges toward the product of the* $m \cdot s$ *marginal distributions of the single genes in the initial generation.* This holds under random mating and chromosome segregation.

From this theorem follows a result for the d.zy. by means of equation (4) of I.

5. *The General Limit Theorem.*

We still have to consider the various possibilities connected with a "degenerated" s.d. The "general case" studied in the preceding section gave the result:

$$p^{(n)}(a_{11} \cdots\cdots a_{1s} \cdots\cdots a_{m1} \cdots\cdots a_{ms}) \to p(a_{11}) \cdots\cdots p(a_{ms}), \tag{43}$$

if, and only if, *no* exchange probability (and consequently no recombination probability) is zero. If, on the other hand, *all* r.p.'s (and consequently all exchange probabilities) are zero then:

$$p^{(n)}(a_{11} \cdots\cdots a_{ms}) = p(a_{11} \cdots\cdots a_{ms}), \quad (n = 0, 1, \cdots\cdots). \tag{44}$$

In these last formulas we again used the notation $p^{(0)} = p$. Between these two extreme cases are the various cases of groups of linked genes.

Let us first consider an example which will exhibit the various possibilities. Assume $m = 4$, $s = 3$, $c_1 = c_2 = 0$, $c_3 \neq 0$, $c_4 \neq 0$, and all six r.p.'s different from zero. Avoiding double subscripts we first consider the marginal distribution of the three genes at the first locus.

The recurrence formula then reads

$$p^{(n+1)}(a_1a_2a_3) = 2l_1(0)p^{(n)}(a_1a_2a_3) + \cdots .$$

Since $c_1 = 0$, i.e., $2l_1(0) = 1$, it follows that

$$p^{(n)}(a_1a_2a_3) = p(a_1a_2a_3), \quad (n=0, 1, \cdots),$$

which means these three alleles are completely linked. In the same way it is seen that, because of $c_2 = 0$:

$$p^{(n)}(b_1b_2b_3) = p(b_1b_2b_3), \quad (n=0, 1, \cdots).$$

On the other hand consider the recurrence formula for

$$p^{(n+1)}(a_1a_2a_3b_1b_2b_3).$$

The coefficients at the right-hand side of this recurrence formula are the values of $l_{12}^{(0)(0)}(\varepsilon_1\ \varepsilon_2)$, $(\varepsilon_i = 0, \cdots 3)$. Let us write for the moment $l(\varepsilon_1\ \varepsilon_2)$ instead of $l_{12}^{(0)(0)}(\varepsilon_1\ \varepsilon_2)$. There are 16 such terms. Since $l_1(1) = 0$, $l_2(1) = 0$, and the corresponding symmetry relations hold, it follows that of these 16 terms 12 are zero and $l(03) = l(30)$, $l(00) = l(33)$ are the only ones not necessarily zero. Next we see that $l(03) = l(30) \neq 0$ because, as can be easily seen, c_{12} would otherwise vanish. If $l(00) \neq 0$ we find, as usual, by means of Lemma A) that

$$p^{(n)}(a_1a_2a_3b_1b_2b_3) \to p(a_1a_2a_3)p(b_1b_2b_3),$$

and, if $l(00) = 0$, it follows that for all n:

$$p^{(n)}(a_1a_2a_3b_1b_2b_3) = p(a_1a_2a_3)p(b_1b_2b_3).$$

On the other hand, since $c_3 \neq 0$, we see easily that

$$p^{(n)}(c_1c_2c_3) \to p(c_1)p(c_2)p(c_3),$$

and similarly for the fourth locus.

If, finally, we consider the joint distribution of all four loci it may be seen by analogous considerations that

$$\lim_{n\to\infty} p^{(n)}(a_1a_2a_3 \cdots d_1d_2d_3) = p(a_1a_2a_3)p(b_1b_2b_3)p(c_1) \cdots p(d_3) \cdot$$

As a second example consider the same situation but with $c_{12} = 0$. It then follows for the linkage distribution $l(\varepsilon_1\ \varepsilon_2) = l_{12}^{(2)(2)}(\varepsilon_1\ \varepsilon_2)$ that $l(00) = 1/2$, and consequently

$$p^{(n)}(a_1a_2a_3b_1b_2b_3) = p(a_1a_2a_3b_1b_2b_3), \quad (n=0, 1, \cdots),$$

and for the joint distribution

$$\lim_{n\to\infty} p^{(n)}(a_1 \cdots d_3) = p(a_1 \cdots b_3)p(c_1) \cdots p(d_3).$$

If, next, not only c_{12} but also c_{13} is equal to zero it follows from the "triangular relation" for the r.p.'s (see Geiringer, 1944, page 34) that $c_{23} = 0$, and from Lemma D) that $c_1 = c_2 = c_3 = 0$. Assume $c_4 > 0$, then necessarily $c_{14} > 0$, $c_{24} > 0$, $c_{34} > 0$, and we obtain

$$\lim_{n\to\infty} p^{(n)}(a_1 \cdots\cdot d_3) = p(a_1 \cdots\cdot c_3)\, p(d_1)\, p(d_2)\, p(d_3)\,.$$

In fact, according to the above results we may say:

If $c_i = 0$ *there is complete linkage between the* s *genes which correspond to the* ith *locus. If* $c_{ij} = 0$, *and consequently* $c_i = c_j = 0$, *there is complete linkage between the* 2s *genes which correspond to these two loci.*

If we want to completely describe the various possibilities we have to introduce the concept of *maximal groups of completely linked factors,* used first in Geiringer (1945). Consider the m numbers $1, 2, \cdots\cdot m$. There are $t \geqq 0$ disjoint groups $G_1, G_2, \cdots\cdot G_t$ each containing $u_i \geqq 1$ of the m numbers $1, 2, \cdots\cdot m$ $(i = 1, \cdots\cdot t)$ such that *within* each group all r.p.'s are zero but between any two such groups at least one r.p. is different from zero; this holds if $u_i \geqq 2$. In addition we speak of a "group of size one" $(u_i=1)$ if the corresponding $c_i=0$ (the exchange probabilities which belong to numbers contained in a group of size > 1, are a fortiori zero). The remaining numbers which do not belong to any "group" and for which therefore the corresponding $c_j > 0$, we call "free elements." We may describe the situation as follows: Within each group no recombination takes place, but there is at least one r.p. different from zero between any two Groups G_i and G_h, and between any "free" element and any other number (whether "free" or in a "group"). If, for example, for $m=7$ we know that $c_3 \neq 0$, $c_5 \neq 0$, $c_{12} = c_{14} = c_{46} = c_{47} = 0$, the triangular relation $c_{67} \leqq c_{46} + c_{47}$ shows that $c_{67} = 0$; hence $(4, 6, 7)$ form a group, but this is not a "maximal" group. Moreover $(1, 2)$ form a group, since $c_{12} = 0$; but in addition $c_{14} = 0$; thence, because of $c_{24} \leqq c_{12} + c_{14}$, $c_{24} = 0$, and continuing that way we see that $c_{14} = c_{16} = c_{17} = c_{24} = c_{26} = c_{27} = 0$ and the five elements $(1,2,4,6,7)$ form a maximal group of completely linked factors while "3" and "5" are "free elements."

Now assume that $G_1, \cdots\cdot G_t$ are the disjoint maximal groups of completely linked factors. Then for the gametic probability correresponding to each such group (marginal distribution of order $s \cdot u_i$) we have

$$p_{G_i}^{(n)} = p_{G_i} \quad (i=1, \cdots\cdot t),\ (n=0,1, \cdots\cdot)\,, \tag{45}$$

which means the complex of $s \cdot u_i$ genes corresponding to G_i forms a group of complete linkage and is conserved through the generations. This follows in the usual way by writing down the recurrence formula for $p_{G_i}^{(n+1)}$ where on the right-hand side the coefficient of $p_{G_i}^{(n)}$ equals one. Next consider the joint (marginal) distribution of two groups G_1 and G_2, namely $p^{(n)}{}_{G_1 G_2}$. We see in the same way as in the examples, using the fact that no r.p. between the numbers of G_1 and the numbers of G_2 equals zero, that the recurrence formula is of the form

$$p^{(n+1)}{}_{G_1 G_2} = \alpha\, p^{(n)}{}_{G_1 G_2} + \beta\, p^{(n)}{}_{G_1}\, p^{(n)}{}_{G_2}$$

with $\alpha + \beta = 1$ and $\beta > 0$, and consequently either

$$p^{(n)}{}_{G_1 G_2} = p_{G_1}\, p_{G_2}, \quad (n=0,1,\cdots), \quad \text{or} \quad p^{(n)}{}_{G_1 G_2} \to p_{G_1}\, p_{G_2}.$$

Hence certainly $\lim_{n\to\infty} p^{(n)}{}_{G_1 G_2} = p_{G_1}\, p_{G_2}$. Following up these ideas we see that

$$\lim_{n\to\infty} p^{(n)}{}_{G_1 G_2 \cdots G_t} = p_{G_1}\, p_{G_2} \cdots p_{G_t}.$$

On the other hand the s genes which correspond to a free element are not linked since the corresponding exchange probability is different from zero. If j is the number of such an element we have

$$\lim_{n\to\infty} p^{(n)}(a_{j_1} \cdots a_{j_s}) = p(a_{j_1}) \cdots p(a_{j_s}).$$

We may formulate these facts in the following

Theorem II. For a 2s-ploid *of* r *alleles and* m *loci consider the heredity assumptions of this paper* (random mating, chromosome segregation, etc.). *Denote by* $G_1, \cdots G_t$ $(0 \leqq t \leqq m)$ *the "maximal groups of completely linked genes," each group containing* $u_i \geqq 1$ *numbers* $(i=1, \cdots t)$, $(u_1 + \cdots + u_t = u \leqq m)$, *and by* j , ···· , k *the remaining* f *free elements* $(0 \leqq f \leqq m-u)$ *with non-vanishing exchange probabilities. Then:*

1) *The complex of* $(s \cdot u_i)$ *genes corresponding to each group* G_i *remains permanently linked, while with respect to the free elements the distribution of each single gene only is preserved*

$$p^{(n)}{}_{G_1} = p_{G_1}, \cdots p^{(n)}{}_{G_t} = p_{G_t} \quad (n=0,1,\cdots); \tag{46}$$

$$p^{(n)}(a_{j\sigma}) = p(a_{j\sigma}), \cdots\ p^{(n)}(a_{k\sigma}) = p(a_{k\sigma}); \quad (n=0,1,\cdots), (\sigma=1,\cdots s). \tag{47}$$

2) *The joint distribution of all* m · s *genes converges toward a limit such that the original marginal probabilities corresponding to each maximal group as well as the* f · s *original marginal probabilities of the remaining single genes are independently distributed.*

If $f=m$, which means if there are free elements only, our theorem reduces to that of the preceding section. For $s=1$, it gives the general result of Geiringer (1945). (If $s=1$ the new concept of a "group" of size one with vanishing exchange probability has no realization.) If $m=1$ and general s the exchange probability, c, with respect to the one locus is either zero or not and accordingly (41) or (43) of Geiringer (1948a) holds.

As an example consider $m=9$, $s=4$, and assume we know that $c_{12}=c_{13}=c_{14}=0$, $c_{56}=c_{57}=0$, $c_8=0$ but $c_9\neq 0$, $c_{15}\neq 0$, $c_{18}\neq 0$. It follows that $G_1=(1,2,3,4)$, $G_2=(5,6,7)$, $G_3=(8)$ are the three maximal groups and "9" is a free element. We then have:

$$\lim_{n\to\infty} p^{(n)}(a_{11}\cdots\cdot a_{94}) = p(a_{11}\cdots\cdot a_{44})\; p(a_{51}\cdots\cdot a_{74}) \cdot p(a_{81}\cdots\cdot a_{84})\; p(a_{91})\cdots\cdot p(a_{94}).$$

In formulating our main result we may as well start with the consideration of the free elements and determine the t disjoint maximal groups of complete linkage for the remaining $(m—f)$ numbers. This gives:

Theorem III. Denote by $c_1, c_2, \cdots\cdot c_f$ *the non vanishing exchange probabilities* $(0 \leqq f \leqq m)$, *and by* $G_1, G_2, \cdots\cdot G_t$ *the maximal groups of completely linked genes, each containing* $u_i \geqq 1$ *of the remaining* (m—f) *numbers* $(u_1+\cdots\cdot u_t = m-f)$. *Then under the hypotheses of this paper: a) The original (marginal) probability of each of the* (m · s) *single genes is preserved through the generations. b) Within each group the original (marginal) probability of the whole group is preserved. c) The joint gametic distribution* (m · s*-variate distribution) in successive generations approaches a limit where genes and groups of genes are independently distributed.*

From these theorems we derive corresponding limit results for the distributions of zygotes by means of equation (4) of I.

LITERATURE

De Winton, D. and Haldane, J. B. S. 1931. "Linkage in the Tetraploid Primula Sinensis." *Jour. Genetics*, **24**, 121–144.

Fisher, R. A. 1947. "The Theory of Linkage in Polysomic Inheritance." *Phil. Trans. Roy. Soc. London*, **B., 233**, 55–87.

Geiringer, H. 1944. "On the Probability Theory of Linkage in Mendelian Heredity." *Ann. Math. Stat.*, 15, 25–57.

Geiringer, H. 1945. "Further Remarks on Linkage in Mendelian Heredity." *Ann. Math. Stat.*, **16**, 390–398.

Geiringer, H. 1948. "Contribution to the Heredity Theory of Multivalents." *Jour. Math. and Phys.*, **26**, 246–278.

Haldane, J. B. S. 1930. "Theoretical Genetics of Autopolyploids." *Jour. Genetics*, **22**, 359–372.

Jennings, H. S. 1917. "The Numerical Results of Diverse Systems of Breeding with Respect to Two Pairs of Characters, etc." *Genetics*, **2**, 97–154.

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 11, 1949

A MECHANISM FOR ACTIVE TRANSPORT OF COENZYME-LIKE SUBSTANCES WITH POSSIBLE REFERENCE TO AUXIN TRANSPORT IN PLANTS

J. BOYD BEST, JR.
COMMITTEE ON MATHEMATICAL BIOLOGY
THE UNIVERSITY OF CHICAGO

A geometrically constrained enzyme system is presented which is capable of driving a substance, which acts as a coenzyme to one of the enzymes of the system, against a concentration gradient. The energy for the process would be derived from the breakdown of the substrate for the enzyme-coenzyme system. If auxin-like substances act in the manner of a coenzyme in plants, then the mechanism presented might serve as a model to explain the polar transport of such substances in plants.

In the following paper a mechanism is outlined which could transport polarly, against a concentration gradient, a substance that acts as a coenzyme to some enzyme constrained in the transporting cells. The energy for the work done against the concentration gradient would be derived from the breakdown of the substrate of the enzyme-coenzyme system.

If such growth stimulating substances as indole-3-acetic acid and the so-called "auxins" in plants act as coenzymes to some enzyme fixed in the plant cells (Commoner and Thimann, 1941; Sweeney and Thimann, 1942; Bose, Dutt and Thakurta, 1944-1946), then the mechanism to be described might serve as a theoretical model to explain the polar transport of these substances in plant coleoptiles. Although this theoretical model was conceived with reference to the problem of polar transport of "growth hormones" in plant coleoptiles, it should be pointed out that its possible application is by no means restricted to this particular case.

This type of mechanism bears a similarity to a more general one, which has been discussed previously by J. Franck and J. E. Mayer (1937) without any reference to this particular problem, and also to the mechanism which has been proposed for active uptake of glucose into the gut by means of phosphorylation and dephosphorylation (Höber, 1945).

In the present theory, the coenzyme-like substance, which for brevity shall simply be designated as "*I*," is transported the length of each cell by protoplasmic streaming, but transported polarly across the boundary regions from one cell to the next by a "pumping" mechanism to be discussed presently. The protoplasm at the end regions is, in this model, to possess the mechanism for polar transport.

If the model is to be applied to the specific problem of auxin transport, then the transporting cells, indicated schematically in Figure 1, would be a series of parenchymal cells of the coleoptile.

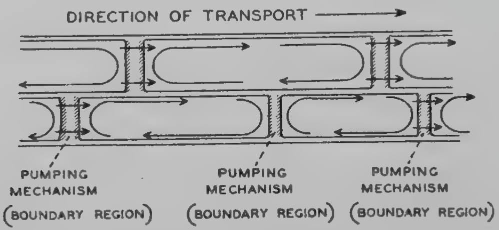

DIRECTION OF TRANSPORT
PUMPING MECHANISM (BOUNDARY REGION)
PUMPING MECHANISM (BOUNDARY REGION)
PUMPING MECHANISM (BOUNDARY REGION)

FIGURE 1

The streaming protoplasm acts in the manner of a "conveyer belt" and the pumping mechanism of the end regions serves to unload *I* from the conveyer of one cell and load it onto the conveyor of the next cell.

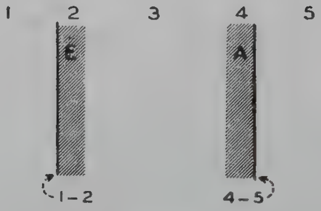


FIGURE 2

Figure 2 shows a highly magnified diagram of the boundary region, whose relation to the rest of the cell is indicated by Figure 1. It is this mechanism, whose geometry is illustrated in Figure 2 (a discussion of the kinetics will occupy the major portion of the remaining part of the paper), which is conceived as driving *I* against the concentration gradient.

Let us consider the mechanism of Figure 2. Suppose the protoplasmic layer comprising this boundary region is divided into three layers—2, 3, and 4. Let the entire system be bounded on both sides by membranes (1-2) between region 1 and the layer 2 and (4-5) between layer 4 and region 5. Layer 2 is distinguished from the other layers of the system by the fact that it contains imbedded in it an enzyme, *E* , for which *I* is the coenzyme. The protoplasm between the two membranes which bound the system is visualized as being gel-like in its structure so that the enzyme, *E* , is geometrically constrained to layer 2. Layer 3 is neutral insofar as the particular set of reactions under consideration is concerned and enters into the present problem solely as an impedance to diffusing substances. Although layer 3 is treated as merely a homogeneous gel in this discussion, in the plant it would probably consist of a relatively impermeable lignin and cellulose wall pierced here and there by threads of protoplasmic gel through which lighter molecular weight substances could diffuse from one cell to the next. As long as layer 3 is not completely impermeable, however, its precise internal structure is of no concern to the present problem. Layer 4 contains, imbedded in its gel, an enzyme A .

One of the simplest and most common types of general reactions involving an enzyme E and a coenzyme I with some substrate S is:

$$\begin{aligned} &E + I \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} EI; \\ &EI + S \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} EIS; \\ &EIS \xrightarrow{k_3} E + IB + \text{Other products.} \end{aligned} \tag{1}$$

The substance IB then reacts with A , the complex breaking down to release I and A again:

$$\begin{aligned} &A + IB \rightleftarrows AIB; \\ &AIB \rightarrow A + I + B . \end{aligned} \tag{2}$$

In our model, reactions (1) take place in layer 2, IB diffuses from layer 2 to layer 4 where reactions (2) occur, releasing I . Since I is consumed in layer 2, the concentration of I there will normally be less than in 1 and also less than in that part of region 3 immediately adjacent to 2. Thus I will diffuse into layer 2 from both 1 and 3. The quantity I is being released in layer 4, however, so that the concentration of I in 4 will be higher than in the immediately adjacent regions of both 3 and 5, provided certain conditions, which will be discussed later, are met. From the preceding considerations of a qualitative nature it is apparent that part of I will diffuse back across layer 3 to 2 and part of it will diffuse into 5.

When the system is operating in a steady state, I will be entering from 1 as fast as it is diffusing into 5. The net overall effect is that I is being transported from 1 to 5 and if the concentration of I in 5 is greater than in 1, this will amount to active transport, i.e., I is being transported against a concentration difference. The assumption has been made that the bounding membranes are permeable to I but impermeable to the larger molecules of IB, so that IB can only diffuse through layer 3 from 2, where it is formed, to 4, where it is consumed. It is not absolutely necessary to the working of the model that the membranes be completely impermeable to *IB*, but it simplifies the mathematical analysis to assume complete impermeability to *IB*.

The energy for the pumping action comes from the overall reaction, which is the breakdown of *S* into "other products" and *B*.

In the subsequent discussion, we shall introduce the following notations:

I_1 = concentration of I in 1;
I_2 = average concentration of I in 2;
I_4 = average concentration of I in 4;
I_5 = concentration of I in 5;
δ_2 = thickness of layer 2;
δ_3 = thickness of layer 3;
δ_4 = thickness of layer 4;
IB_2 = average concentration of IB in 2;
IB_4 = average concentration of IB in 4;
v_2 = average velocity of reaction (1) in layer 2;
v_4 = average velocity of reaction (2) in layer 4;
$\Delta = I_5 - I_1$.

Care must be exercised in making the kinetics formulations on the assumption of steady state conditions, because we have said that the last stage of reactions (1) and (2) proceeds irreversibly. If I_2 is made sufficiently large so that

$$\delta_2 v_2 > \delta_4 v_{4\ \max}\,, \tag{3}$$

then a steady state is impossible, since the amount of *IB* in the system will continue to increase without limit. To insure against this defect in the formulation, one of two alternatives is available. First, the last stage of reaction (1) could be made reversible, but then some assumption would have to be made about the nature of the products. Secondly, one could assume conditions such that v_2 was the limiting reaction, i.e., such that

$$\delta_2 v_{2\ \max} < \delta_4 v_{4\ \max}\,. \tag{4}$$

The latter assumption will be chosen because it is less drastic and less restricting than the assumption of some specific nature of the products. More will be said of this after the kinetics formulations have been made.

A flow scheme for the various substances involved in the mechanism is indicated in Figure 3.

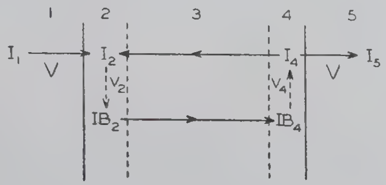


FIGURE 3

The broken arrows labeled v_2 and v_4 indicate chemical reactions while the full arrows indicate directions of diffusion. Having now given a qualitative description of the theoretical model, let us proceed to a quantitative investigation of its properties.

The Kinetics Formulation.

Let k_1, k_{-1}, k_2, k_{-2} and k_3 be the reaction velocity constants as indicated for reactions (1). Then we have for steady state conditions:

$$\begin{aligned} v_2 &= k_1(E)(I) - k_{-1}(EI); \\ 0 &= k_1(E)(I) + k_{-2}(EIS) - k_{-1}(EI) - k_2(EI)(S); \\ 0 &= k_2(EI)(S) - k_{-2}(EIS) - k_3(EIS). \end{aligned} \tag{5}$$

If E_0 is the total amount of enzyme per unit volume in layer 2, then

$$0 = E_0 - E - EI - EIS. \tag{6}$$

Solution of equations (5) and (6) gives

$$v_2 = \frac{(E_0)(I_2)(S)k_1k_2k_3}{k_1k_2(I_2)(S) + k_1(k_{-2}+k_3)(I_2) + k_2k_3(S) + k_{-1}(k_{-2}+k_3)}. \tag{7}$$

Similarly for the reaction (2) occurring in layer 4, we have

$$v_4 = \frac{(A_0)(IB_4)k_5}{(IB_4) + \left(\dfrac{k_{-4}+k_5}{k_4}\right)} = \frac{(IB_4)W}{(IB_4) + K'}, \tag{8}$$

where

$$\begin{aligned} W &= k_5(A_0), \\ K' &= k_{-4} + k_5/k_4. \end{aligned}$$

Let us now see what assumption (4) involves. For a fixed value of S

$$v_2 \rightarrow \frac{E_0 k_2 k_3 S}{k_2 S + k_{-1}(k_{-2} + k_3)} \quad \text{as } I_2 \rightarrow \text{large}, \tag{9}$$

or, if both I_2 and S become large,

$$v_2 \rightarrow E_0 k_3 = v_{2\ \max}. \tag{10}$$

Then the condition under which the steady state formulation applies is that

$$E_0 k_3 < \frac{\delta_4}{\delta_2} k_5 A_0 .$$

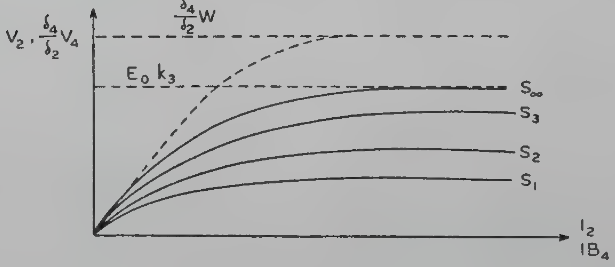


FIGURE 4

Figure 4 illustrates graphically v_2 as a function of I_2 for various concentrations of S (full lines). On the same set of axes and to the same scale is also shown

$$\frac{\delta_4}{\delta_2} v_4$$

as a function of IB_4 (broken line curve).

From assumption (4) it follows that $E_0 k_3$ must always be less than

$$\frac{\delta_4}{\delta_2} W,$$

and so, since for steady state

$$v_2 < \frac{\delta_4}{\delta_2} v_4,$$

the system will usually operate in the region where

$$\frac{\delta_4}{\delta_2} v_4$$

is approximately linear. This approximation does not enter into or affect the final result since v_2 completely fixes v_4, but it does somewhat simplify the algebra involved in the elimination of v_4 from our equations.

Diffusion Relations and Boundary Conditions for Steady State.

Let D_1 = diffusion coefficient of I through 3;
D_2 = diffusion coefficient of IB through 3;
V = flux of I through the system (amount of transport), expressed in gm cm^{-2} sec^{-1};

h_2 = permeability of membrane 1-2 to I;
h_4 = permeability of membrane 4-5 to I.

From condition of steady state, the rate of consumption of I by reaction (1) must be the rate at which I diffuses into layer 2, so that

$$\delta_2 v_2 = h_2(I_1 - I_2) + D_1 \frac{I_4 - I_2}{\delta_3}. \tag{11}$$

Similarly, the rate at which I is liberated by reaction (2) in layer 4 must be equal to the rate of diffusion of I out of 4, so that

$$\delta_4 v_4 = h_4(I_4 - I_5) + D_1 \frac{I_4 - I_2}{\delta_3}. \tag{12}$$

Since the membrane 1-2 is impermeable to IB, the rate of production of IB by reaction (1) is equal to the rate of diffusion of IB across layer 3 and the rate of consumption of IB there by reaction (2), hence

$$\delta_2 v_2 = D_2 \frac{IB_2 - IB_4}{\delta_3}; \tag{13}$$

$$\delta_4 v_4 = D_2 \frac{IB_2 - IB_4}{\delta_3}. \tag{14}$$

If the total amount of the various forms of I in the system is to remain constant, then the inflow of I across one of the membranes must be equal to the outflow of I across the other membrane, so that

$$V = h_2(I_1 - I_2). \tag{15}$$

From the set of equations (11)—(15) we obtain, after some algebraic manipulation, the expression

$$v_2 = \frac{D_1}{\delta_2\,\delta_3}\Delta + V\frac{D_1 h_4 + D_1 h_2 + \delta_3 h_2 h_4}{\delta_2\,\delta_3 h_2 h_4}. \tag{16}$$

The form of (16) may be somewhat simplified if we put

$$H = \frac{h_4 + h_2}{h_4 h_2}, \tag{17}$$

so that (16) takes the form

$$v_2 = \frac{D_1 \Delta}{\delta_2 \delta_3} + \frac{V(D_1 H + \delta_3)}{\delta_2 \delta_3} = \alpha. \tag{18}$$

The Solution for V.

If we solve (15) for I_2 explicitly and substitute this into the kinetics formulation (7), putting

$$K = \frac{k_{-2} + k_3}{k_2},$$

we find

$$v_2 = \frac{k_3 E_0 S \left(I_1 - \frac{1}{h_2} V \right)}{(S + K) \left(I_1 - \frac{1}{h_2} V \right) + \frac{k_3}{k_1} S + \frac{k_{-1}}{k_1} K} = \beta. \tag{19}$$

For steady state operation, the two expressions (18) and (19) for v_2 must be satisfied simultaneously.

For convenience in manipulation and to make more apparent the form of the relations, put

$$M = \frac{D_1}{\delta_2 \delta_3}; \tag{20}$$

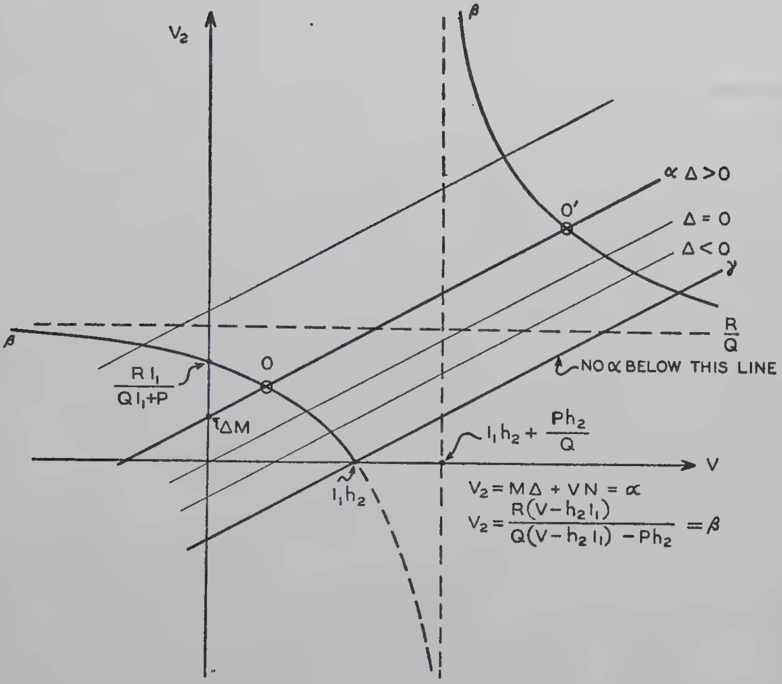
$$N = \frac{D_1 H + \delta_3}{\delta_2 \delta_3}; \tag{21}$$

$$Q = (S + K); \tag{22}$$

$$P = \frac{k_3 S + k_{-1} K}{k_1}; \tag{23}$$

$$R = k_3 E_0 S. \tag{24}$$

The behavior of V can best be examined graphically (Figure 5). If v_2 is plotted against V, equation (19) gives rise to a two branched curve, with a different pair of curves for each value of I_1. The vertical asymptote is shifted to the right along the positive direction of



V_2
β
$\alpha \; \Delta > 0$
O'
$\Delta = 0$
$\Delta < 0$
γ
$\frac{R}{Q}$
β
NO α BELOW THIS LINE
$\frac{RI_1}{QI_1+P}$
O
ΔM
$I_1h_2 + \frac{Ph_2}{Q}$
V
I_1h_2
$V_2 = M\Delta + VN = \alpha$
$V_2 = \frac{R(V-h_2I_1)}{Q(V-h_2I_1) - Ph_2} = \beta$

FIGURE 5

the V axis for increasing values of I_1, while the horizontal asymptote remains the same for all values of I_1. The vertical asymptote of the pair of β curves will obviously always lie in the right half of the coordinate plane, since $I_1 \geq 0$ and $Ph_2/Q > 0$.

Equation (18) represents a family of straight lines, α, of constant slope but with different intercepts on the v_2 axis. Each value of Δ gives rise to a different α line. The intersections O and O' of the α and β give the two values of V that arise from simultaneous solution of (18) and (19). It will be shown, however, that only the intersection O has any physical meaning. It will also be shown later that this intersection O represents a stable configuration point for the system.

It should be noted that the straight line γ represents a bound for all the α lines. That is to say, all the α lines must lie above γ. The line γ has the same slope as the α lines and intersects the V axis at I_1h_2. The reason for γ arises from the following physical considerations.

For an α line to cross the V axis at a value larger than I_1h_2, it would be necessary for it to meet the condition that

$$-\frac{M}{N}\Delta > I_1 h_2 , \tag{25}$$

or that

$$\frac{D_1\,\Delta}{D_1 H + \delta_3} > I_1 h_2 , \tag{26}$$

which is tantamount to saying that in the pure diffusion case, i.e., with no driving reaction (see page 234 of this paper), the diffusion across the entire system is greater than the flow across the bounding membrane (1-2). *This is obviously impossible.*

Further information can be obtained from Figure 5. The limiting value of Δ against which the system can pump, for infinitely large I_1 is

$$\Delta_{\max} = \frac{R}{MQ} = \frac{\delta_2\,\delta_3\,k_3 E_0 S}{D_1 (S+K)} . \tag{27}$$

The limiting or maximum transport, for infinitely large I_1, corresponding to any value of Δ is given by

$$\left(\frac{R}{Q} - \Delta M\right)\frac{1}{N} = V_{\max}, \text{ where} -\frac{M}{N}\Delta \leq I_1 h_2 ;$$

or

$$\left[\frac{k_3 E_0 S}{(S+K)} - \Delta\frac{D_1}{\delta_2\,\delta_3}\right]\frac{\delta_2\,\delta_3}{(D_1 H + \delta_3)} = V_{\max} . \tag{28}$$

Analysis of the Non-Steady State.

Up to now in our investigation of the theoretical model, which has been postulated, the steady state has been assumed in all of the formulations. Two roots in the solution for V were obtained, giving two values of V corresponding to a given I_1 and Δ. We shall now study these configuration points with regard to their stability, to ascertain whether they represent stable dynamic equilibria and which root would represent the actual steady state configuration point of the physical system.

For layer 2 the rate of change in the amount of I in the layer is given by the expression

$$\delta_2 \frac{dI_2}{dt} = h_2 (I_1 - I_2) + \frac{D_1}{\delta_3}(I_4 - I_2) - \delta_2 v_2 , \tag{29}$$

and for layer 4 by

$$\delta_4 \frac{dI_4}{dt} = \delta_4 v_4 - h_4 (I_4 - I_5) - \frac{D_1}{\delta_3} (I_4 - I_2) . \tag{30}$$

Since all of the layers will be thin, the approximation will be made that the gradient of *IB* across 3 always adjusts rapidly enough so that we can consider the flow of *IB* into 3 to be equal to the flow of *IB* out of 3. This approximation has also been made for I, in writing expressions (29) and (30). Since there is practically no delay in the adjustment of v_4 to new concentrations of *IB*, we can write as before

$$\delta_2 v_2 = \delta_4 v_4 . \tag{31}$$

Introducing equations (7), (20), (21), (22), (23), (24), and (31) we can now write (29) and (30) in the form

$$\frac{dI_2}{dt} = \frac{h_2}{\delta_2} I_1 - \left[\frac{h_2 \delta_3 + D_1}{\delta_2 \delta_3} + \frac{R}{QI_2 + P} \right] I_2 + \frac{D_1}{\delta_2 \delta_3} I_4 ; \tag{32}$$

$$\frac{dI_4}{dt} = \frac{h_4}{\delta_4} (I_1 + \Delta) + \left(\frac{D_1}{\delta_3 \delta_4} + \frac{\delta_2 R / \delta_4}{QI_2 + P} \right) I_2 - \left(\frac{h_4 \delta_3 + D_1}{\delta_3 \delta_4} \right) I_4 . \tag{33}$$

Then the curve $dI_2/dt = 0$ will be given by

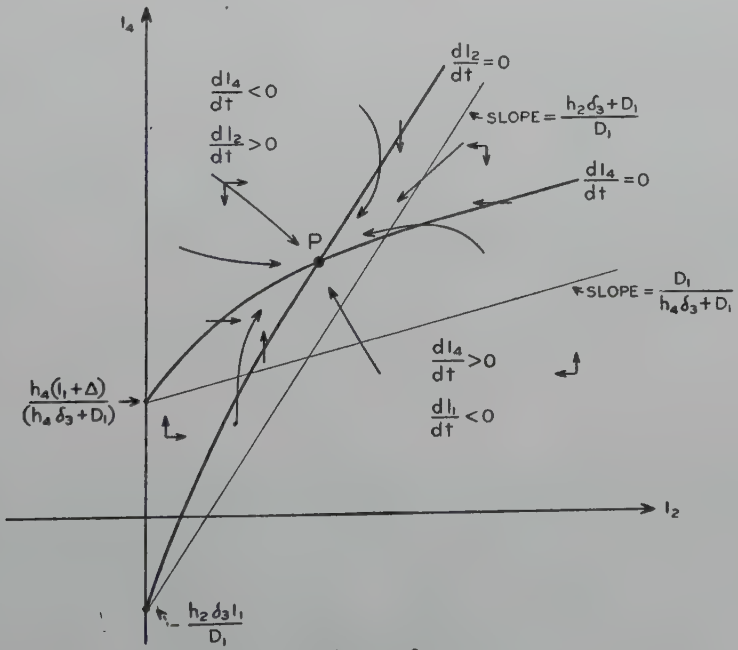

I_4
$\frac{dI_2}{dt} = 0$
$\frac{dI_4}{dt} < 0$
$\frac{dI_2}{dt} > 0$
SLOPE = $\frac{h_2\delta_3 + D_1}{D_1}$
$\frac{dI_4}{dt} = 0$
P
SLOPE = $\frac{D_1}{h_4\delta_3 + D_1}$
$\frac{dI_4}{dt} > 0$
$\frac{dI_1}{dt} < 0$
$\frac{h_4(I_1+\Delta)}{(h_4\delta_3 + D_1)}$
I_2
$\frac{h_2\delta_3 I_1}{D_1}$

FIGURE 6

$$-\frac{h_2\,\delta_3}{D_1}I_1+\left(\frac{h_2\,\delta_3+D_1}{D_1}\right)I_2+\left(\frac{\delta_2\,\delta_3\,R}{D_1}\right)\frac{I_2}{(QI_2+P)}=I_4, \tag{34}$$

and $dI_4/dt=0$ will be given by

$$\left(\frac{D_1}{h_4\,\delta_3+D_1}\right)I_2+\left(\frac{\delta_2\,\delta_3\,R}{h_4\,\delta_3+D_1}\right)\frac{I_2}{(QI_2+P)}+\frac{h_4\,\delta_3\,(I_1+\Delta)}{(h_4\,\delta_3+D_1)}=I_4. \tag{35}$$

If I_4 is plotted as a function of I_2 as in Figure 6, it may be seen that the curves representing (34) and (35) have only one point of intersection for positive I_2 and I_4. That this is true may be verified by noting that the slope of both curves approaches a constant value with increasing I_2 and the slope of (34) is greater than (35) for all positive I_2. Thus there is only one steady state root corresponding to $dI_2/dt=0$ and $dI_4/dt=0$ that is physically significant, i.e., not involving negative concentrations, that arises from the solution of (34) and (35). It may be seen from these equations that for I_4 greater and I_2 less than the values lying on the two curves, dI_4/dt is less than 0 and dI_2/dt is greater than 0. This would mean that any point in that region would move downward and to the right with time. With each point it is possible to associate two vector components, one along the direction of the I_4 axis and one parallel to the I_2 axis. A similar procedure may be carried out by examining the values of the two time derivatives for points of each of the other regions bounded by the two curves. This is done in Figure 6 and the longer arrows indicate the resultant direction of motion of points of the various regions of the I_4 vs. I_2 plane. Figure 6 shows that this point given by the intersection of the two curves is a stable point, since regardless of what the initial values of I_2 and I_4 are they will, with increasing time, approach those values corresponding to the point P.

Solving (34) and (35) analytically to eliminate I_4 and obtain a solution for I_2 in terms of I_1 and Δ, we have

$$I_2=-\tfrac{1}{2}\left(\frac{P+b}{Q}-C\right)\pm\tfrac{1}{2}\sqrt{\left(\frac{P+b}{Q}-C\right)+\frac{4CP}{Q}}, \tag{36}$$

where

$$b=\frac{\delta_2\,\delta_3{}^2\,Rh_4}{(h_2h_4\delta_3{}^2+h_4\delta_3D_1+h_2\delta_3D_1)}, \tag{37}$$

and

$$C=\frac{(\delta_3h_4D_1+\delta_3{}^2h_2h_4+h_2\delta_3D_1)I_1+\delta_3h_4D_1\Delta}{(h_2h_4\delta_3{}^2+h_4\delta_3D_1+h_2\delta_3D_1)}. \tag{38}$$

The constant C is always positive, since when Δ is negative $|\Delta| < I_1$. All other terms in C are positive. The quantity b will always be positive, since it involves only positive quantities. Similarly we have seen previously that P and Q are both positive. Thus

$$\sqrt{\left(\frac{P+b}{Q}-C\right)^2+\frac{4CP}{Q}} > \left(\frac{P+b}{Q}-C\right), \tag{39}$$

so that only the root obtained by choosing the positive sign of the radical will have any physical significance, i.e.,

$$I_2 = -\tfrac{1}{2}\left(\frac{P+b}{Q}-C\right)+\tfrac{1}{2}\sqrt{\left(\frac{P+b}{Q}-C\right)^2+\frac{4CP}{Q}}. \tag{40}$$

Examination of the case when $\Delta = 0$ and equation (38) reduces to $C = I_1$ shows that from (15) and (40) we have

$$\begin{aligned} V &= h_2(I_1 - I_2) \\ &= h_2\left(I_1 - \tfrac{1}{2}I_1 + \frac{P+b}{2Q} - \tfrac{1}{2}\sqrt{I_1^2 + 2\frac{(P-b)}{Q}I_1 + \frac{(P+b)^2}{Q^2}}\right), \end{aligned} \tag{41}$$

and as $I_1 \rightarrow$ large

$$V = h_2\left[I_1 - \tfrac{1}{2}I_1 + \frac{P+b}{2Q} - \tfrac{1}{2}I_1\left(1+\frac{(P-b)}{I_1Q}+\cdots\right)\right];$$

$$V = \frac{h_2 b}{Q} = \frac{1}{N}\frac{R}{Q},$$

which we see corresponds to the intersection O of Figure 5 for these conditions.

If, on the other hand, we choose the root giving a negative value of I_2, we obtain by making $\Delta = 0$ and I_1 large:

$$V = h_2 I_1 + \frac{Ph_2}{Q},$$

which, inspection of Figure 5 will show, corresponds to the operating point O' under these circumstances. Hence we see that O' has no physical significance and the point O of Figure 5 is *the* operating point of the system.

The Pure Diffusion Case.

Since reference has been made earlier in this paper to the degenerate case when there is no reaction, but merely diffusion across the system, it shall be derived at this point for purposes of reference. We can write immediately that

$$h_2(I_1 - I_2) = h_4(I_4 - I_5)\,; \tag{42}$$

$$\frac{D_1}{\delta_3}(I_2 - I_4) = h_4(I_4 - I_5)\,; \tag{43}$$

$$V = h_2(I_1 - I_2)\,; \tag{44}$$

from which is obtained upon elimination of I_4 and I_2 that

$$V = \frac{-D_1\,\Delta}{D_1 H + \delta_3}. \tag{45}$$

This work was aided in part by a grant from the Dr. Wallace C. and Clara A. Abbott Memorial Fund of The University of Chicago.

LITERATURE

Bose, D. M., Dutt, B. K. and Thakurta, A. Guha. 1944-1946. "Studies on the Mechanical Pulsation and Respiration in the Motile Leaflets of Desmodium Gyrans." *Transactions of Bose Research Institute* (Calcutta), 16, 121–154.

Commoner, B. and Thimann, K. V. 1941. "On the Relation between Growth and Respiration in the Avena Colepotile." *Jour. Gen. Phys.*, 24, 279–296.

Franck, James and Mayer, Joseph E. 1947. "An Osmotic Diffusion Pump." *Arch. Biochem.*, 14, 297–313.

Höber, Rudolph. 1945. *Physical Chemistry of Cells and Tissues.* Philadelphia: The Blakiston Company.

Sweeney, Beatrice Marcy and Thimann, K. V. 1942. "The Effect of Auxins on Protoplasmic Streaming: III." *Jour. Gen. Phys.*, 25, (6), 841–854.

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 11, 1949

BOOK REVIEW

JOSEPH SCHILLINGER. *The Mathematical Basis of the Arts.* 1948. x + 696 pp. New York: Philosophical Library. $12.00.

A mathematical theory of art is of interest both to the mathematician and to the mathematical biologist. In his pioneering work the late George Birkhoff (1933) made an attempt to evaluate the effect of different types of symmetry, order and other quantitative elements of a visual or acoustical pattern upon its pleasantness. His approach was a purely formal one. He postulated, in a somewhat intuitive way, that certain geometric or temporal characteristics of a pattern possess a certain arbitrary degree of pleasantness or unpleasantness. Taking then a number of concrete patterns he evaluated, on the basis of his postulates, their relative aesthetic values. The importance of Birkhoff's work is not diminished by the fact that experimental determinations by R. C. Davis (1936), who used the rank order method, of the actual relative aesthetic values of some of Birkhoff's patterns decidedly disagreed with Birkhoff's predictions. These experiments show that some of Birkhoff's postulates are inadequate. The fundamental method of approach remains of interest, and the very fact that R. C. Davis's experimental work was suggested by Birkhoff's theoretical studies gives the latter at least the value of a good working hypothesis.

A different approach is made by the mathematical biologist. He begins with a set of assumptions about the mechanism of the central nervous system, and of what processes in it correspond to a feeling of pleasantness or unpleasantness. On that basis he then develops a theory of perception, in particular of visual patterns. The assumed nervous mechanism indicates what geometrical features of a pattern are pleasant and to what extent. The aesthetic value of a perceived object is thus considered a function not of the object only, but of the perceiving individual. Calculation of relative aesthetic values of some of Birkhoff's polygons, which were experimentally studied by R. C. Davis, led to quite a surprising agreement between theory and experiment. Subsequent theoretical and experimental work further improved this agreement (Rashevsky, 1948).

In view of the above, it is natural that an attractive looking volume of almost 700 pages entitled, "The Mathematical Basis of the Arts," should catch the eye and excite the curiosity of the mathematician and of the mathematical biologist. It seems, however, fairly safe to say that both the mathematician and the mathematical biologist will be sorely disappointed in this book. This in itself, the reviewer hastens to remark, does not imply any reflection on the quality of any book.

Joseph Schillinger's book undertakes a detailed *classification** of different numerical relations which occur in various arts. Possible numerical sequences, ratios, symmetries, etc., are painstakingly classified. However, the reader will not find any mathematical argument, proof or deduction, as usually understood by the pure or applied mathematician. No discussion is given of the possible quantitative connection of the different numerical relations with the aesthetic value of a work of art. In no sense does the book present a mathematical theory which could be tested experimentally and which would explain anything.

---

*Italics by the reviewer.

Such a classificatory study may be very valuable, perhaps with the view of being used as a sort of empirical material for later theoretical studies. It would have been, however, more advisable to call the book, "Numerical Basis of the Arts," or, still better, "Numerical Relations in the Arts." The prevailing layman's opinion notwithstanding, *numerical* and *mathematical* are *not* synonymous!

We shall illustrate the above statements only by two examples, which are rather typical of the whole book.

On page 64 the author begins an enumeration of different "Art Forms." Thus Art Form Number One has two aspects: (a) *art of audible sound** (music); and (b) *art of audible word** (poetry). Art Form Number Two is the *"art of touchable mass."** Number Four is "smellable odor." In this way eighteen "Forms" are established.

Then on page 75 we read that, "A complex homogeneous form may consist of combinations by two or three elements . . . . ," and the possible combinations are tabulated. The tabulation takes 3 pages, of which the following is a sample:

"By TWO

5 + 8; 5 + 9; 5 + 10; 5 + 11; 5 + 12; 5 + 13; 5 + 14; 5 + 15;
5 + 16; 5 + 17; 5 + 18.
. . . ."

We see 3 pages of numerals, but hardly anyone would call this mathematics!

Another example may be found on page 344, where more than a page is used to list the 34 particular numerical cases of the expression

$$t = \frac{180^\circ (T - 2)}{T},$$

when T has integer values from 3 to 36.

Apart from the suggested change in the title, the above does not constitute a criticism of the book. The reviewer feels, however, that he cannot avoid making some criticism. He does this with the greatest reluctance because the book is a posthumous work, and therefore the author cannot answer any criticism with which he may have disagreed.

After some interesting preliminary discussion, the first "mathematics" appears on page 31. We quote here a whole paragraph.

"An art structure in its formation (the process of being created) may be expressed as a series of sub-structures in their sequent development and accumulation.

$$\left.\begin{matrix} N \\ S \\ I \end{matrix}\right\} = S_{\mathrm{I}} t_1 + (S_{\mathrm{I}} \div S_{\mathrm{II}}) t_2 + (S_{\mathrm{I}} \div S_{\mathrm{II}} \div S_{\mathrm{III}}) t_3 + \cdots$$
$$+ (S_{\mathrm{I}} \div S_{\mathrm{II}} \div S_{\mathrm{III}} \div \cdots \div S_N) t_N .$$

$S =$ substructure

$t_1, t_2, t_3 =$ consecutive moments in time."

Nowhere in the text is the meaning of the other symbols explained. On page 684, in the glossary, compiled by Arnold Shaw, we read: "S used to denote structural group in *Positional Rotation** . . . . S also used to denote *Synchronization,* or coordination of duration, or *Symmetrization,* the coordination of extensions, as in $S(a:b)$." This does not seem to be of much help! After some guessing, the reviewer is inclined to think that the left-hand side of the above expression means a summation from 1 to N, and should, in conventional terms, read $\sum_{1}^{N}$. The meaning of the symbol $\div$ is left for the reader to guess. If this is a division sign, then what does it mean to divide one substructure by another? For that matter, what is the meaning of the product of a "structure" with time? The paragraph quoted above is followed by pages of purely verbal discussion.

---

*Italics by the author.

A mathematician is likely to find such a loose use of symbols rather irritating. As a further example we find the following on page 85. "A natural integer series is an infinite progression of all the integer numbers between one and any integer number approaching infinity. $\sum_{1}^{n} = 1, 2, 3, \cdots\cdot n$." Apparently in this case the symbol $\sum$ does not mean a summation; or does it? Again the glossary on page 684 fails to elucidate the point. Also on page 85 we read: "*Continuity** is a finite portion of continuum" (*sic*!). Many more such peculiarities are found throughout the book.

The value and power of mathematics lies in its ability to provide general expressions, which make it unnecessary to enumerate all particular cases. From this angle again we come to the conclusion that Schillinger's book can hardly be called mathematical, although it contains page after page of numerical material.

As we have already remarked, such a classificatory work may well prove to be useful in the future, not as a theory, which it decidedly is *not,* but as a sort of semi-empirical numerical material. Whether it will prove to be such or not, the future only will show. The amount of work done by the author in writing this book was tremendous. Possibly the author had some theoretical or practical applications in mind for the future. It is certainly to be regretted that his untimely death leaves that conjecture unanswered forever.

N. RASHEVSKY
Committee on Mathematical Biology
The University of Chicago

*Italics by the author.

LITERATURE

Birkhoff, G. 1933. *Aesthetic Measure.* Cambridge: Harvard University Press.

Davis, R. C. 1936. "An Evaluation and Test of Birkhoff's Aesthetic Measure Formula." *Jour. Gen. Psychol.,* 15, 231–240.

Rashevsky, N. 1948. *Mathematical Biophysics.* Revised Edition. Chicago: University of Chicago Press.